The Bell System Technical Journal
Vol. 55, No. 8, October 1976
Printed in U.S.A.

Distinguishing Stable Probability Measures Part I: Discrete Time

By B. W. STUCK

(Manuscript received April 1, 1976)

A sequence of N, independent, identically distributed, random variables is observed from one of two stable distributions with known parameters. The likelihood-ratio test for discriminating between these two distributions is found explicitly and performance limitations are determined.

When the two distributions differ only in location, the likelihood-ratio test is sensitive to whether the distribution is nongaussian stable ($0 < \alpha < 2$) when nonlinear soft limiting of large deviations is used, or gaussian stable ($\alpha = 2$) when linear processing is used.

When the two distributions differ only in scale, the likelihood-ratio test is sensitive to whether $0 < \alpha < 2$ when nonlinear soft limiting of large deviations is used, or gaussian ($\alpha = 2$) when a chi-squared test is used.

The analysis of the two remaining cases, distinguishing between one of two characteristic indices, and between one of two skewness parameters, parallels the analysis of distinguishing between one of two scale parameters and is only touched upon briefly.

I. INTRODUCTION

The problem of classifying a series of observations as coming from one of two or more possible classes or bypotheses bas received a great deal of attention in the statistical and engineering literature. In many pbysical situations, a variety of disturbances corrupt the observations; rather tban model eacb disturbance separately, it is often argued on pbysical grounds that tbe disturbances add and are independent, and the central limit theorem is invoked to model this sum using a gaussian distribution. This approach is adequate as long as tbe sum is not dominated by one or a few of tbe summands; if one or a few of tbe summands does dominate tbe sum, the disturbances can possibly be modeled as a stable distribution, one member of a family of probability distributions whicb includes the gaussian, by invoking a frequently overlooked generalization of tbe central limit theorem.

The gaussian distribution has enjoyed great popularity in hypothesis testing because it is analytically tractable and because it is the only stable distribution with finite variance. Although it may be argued that mathematical models with infinite variance are physically inappropriate, this view conveniently overlooks the fact that the gaussian distribution is unbounded, which is also a physically inappropriate mathematical model. The gaussian model may adequately model disturbances over a narrow range of amplitudes; an infinite-variance, stable-distribution model may adequately model disturbances over a larger range of amplitudes. Both distributions may be physically inappropriate mathematical models, but the infinite-variance distribution may, in this sense, be the better model. This paper examines several stable-distribution hypothesis-testing problems.

The primary motivation for this work on stable probability measures is drawn from a recent statistical analysis[1] of noise on various telephone lines. This analysis indicated telephone noise may be adequately modeled (on the lines examined) by a sum of sinusoids at various frequencies plus a purely nondeterministic random process that is well characterized by a stable distribution (either gaussian or nongaussian stable). Since only a small number of lines were examined, this analysis is preliminary, awaiting other independent investigations.*

Indirect motivation for this work is drawn from detecting electromagnetic signals at frequencies of 100 kHz or less. Noise at these frequencies is claimed to be nongaussian; unfortunately, adequate statistical evidence to substantiate this claim is lacking, with one exception.[2]

A final source of motivation is found in financial problems. Over the last decade, a large body of statistical evidence has been amassed which indicates that the differences of logarithms of successive equally spaced prices of common stocks can be adequately modeled using stable distributions.[3,4]

II. OUTLINE OF DISCUSSION†

A sequence of N random variables is observed; for simplicity, it is assumed they are independent and identically distributed—drawn from one of two stable distributions with known parameters (characteristic index $0 < \alpha^j \leqq 2$, skewness parameter $-1 \leqq \beta^j \leqq 1$, scale parameter $\gamma^j > 0$, location parameter $-\infty < \delta^j < \infty$; $j = 0, 1$).‡ It

* Applications of this work to removing telephone noise will be presented elsewhere.

† These results were first presented at the Eighth Annual Princeton Conference on Information Sciences and Systems, March 28–29, 1974, p. 405, and at the 1975 Johns Hopkins Conference on Information Sciences and Systems, April 2–4, 1975, pp. 49–51.

‡ Both subscripts and superscripts will be used to denote the stable-distribution parameters under hypothesis $H_j (j = 0, 1)$; these parameters will be discussed more fully in Section III.

is well known that the likelihood-ratio test is a decision rule that is optimum with respect to either a Neyman-Pearson or Bayes criterion.[5] Here, the likelihood ratio is found explicitly and performance limitations of the test are determined. The extension of these results from two to M stable distributions is well known and will not be dealt with here.[5]

The (log) likelihood decision rule, because of the independence assumption, takes the following simple form:

$$\Lambda' = \sum_{i=1}^{N} l(r_i) \underset{H_0}{\overset{H_1}{\gtrless}} L'$$

$$l(r_i) = \ln \frac{p(r_i;\alpha^1,\beta^1,\gamma^1,\delta^1)}{p(r_i;\alpha^0,\beta^0,\gamma^0,\delta^0)},$$

where $\{r_i\}_1^N$ are the N observed random variables, drawn from a distribution with probability density $p(x;\alpha^j,\beta^j,\gamma^j,\delta^j)$, and L' is a threshold. Since $l(r_i)$ can be rewritten as the sum of four functions,

$$l(r_i) = \ln \frac{p(r_i;\alpha^1,\beta^1,\gamma^1,\delta^1)}{p(r_i;\alpha^0,\beta^1,\gamma^1,\delta^1)} + \ln \frac{p(r_i;\alpha^0,\beta^1,\gamma^1,\delta^1)}{p(r_i;\alpha^0,\beta^0,\gamma^1,\delta^1)} + \ln \frac{p(r_i;\alpha^0,\beta^0,\gamma^1,\delta^1)}{p(r_i;\alpha^0,\beta^0,\gamma^0,\delta^1)} + \ln \frac{p(r_i;\alpha^0,\beta^0,\gamma^0,\delta^1)}{p(r_i;\alpha^0,\beta^0,\gamma^0,\delta^0)},$$

each of which tests for only one different parameter, this suggests studying each of these four situations separately.

Two special cases are examined in detail: when the distributions differ only in location and when they differ only in scale. The probabilities of error of the first and second kind are found for three analytically tractable cases (gaussian, Cauchy, and Pearson V) by calculating the characteristic function of the log likelihood probability measure induced under each hypothesis; the general case is apparently analytically intractable, and quite expensive to tackle numerically at present. Exponentially sharp upper and lower bounds on both types of probabilities of error, and also the total probability of error, can be simply derived from the Laplace transform of the log likelihood probability measure induced under each hypothesis. These bounds are found analytically in three cases, and relatively inexpensive numerical results are presented for selected other cases.

When the two distributions differ only in location, the likelihood-ratio test is shown to be extremely sensitive to whether the distribution is nongaussian stable ($0 < \alpha < 2$), when nonlinear soft limiting of large deviations is employed, or gaussian ($\alpha = 2$), when linear processing is used. When the distribution is nongaussian stable, performance is found analytically to be quite sensitive to whether a linear (suboptimum) or likelihood (optimum) decision rule is used: the total

probability of error for tbe linear test behaves asymptotically ($N \gg 1$) as $O(AN^{1-\alpha})$, while the total probability of error for the likelihood-ratio test is upper bounded by $\exp(-BN + C)$, where (A, $B > 0$, C) depend on parameters of the two distributions and are independent of N. (For related work that complements the results in our discussion, see tbe list of references and particularly Refs. 6, 7, and 8.)

When the two distrihutions differ only in scale, the likelihood-ratio test is extremely sensitive to whether the distribution is nongaussian stahle when nonlinear soft limiting of large deviations is used, or gaussian when a chi-squared test is used. Performance for nongaussian stable distributions is extremely sensitive to whether a suboptimum (chi-squared) or optimum (likelihood-ratio) test is used: the total probability of error for the cbi-squared test hehaves asymptotically ($N \gg 1$) as $O(FN^{-(\alpha/2-1)})$, while the total probahility of error for the likelibood-ratio test is upper bounded by $\exp(-GN + H)$, where (F, $G > 0$, H) depend on parameters of the two distributions and are independent of N.

Tbe analysis of the two remaining cases, distinguishing between one of two cbaracteristic indices and between one of two skewness parameters, closely parallels the analysis that distinguishes between two scale factors and is only touched upon here.

The continuous time analogs of these discrete-time problems are studied, where a sample function from one of two stable, stationary, independent-increment processes is ohserved for a finite time interval in the second part of this work. In contrast with this work, the analysis is simpler, and it is possible to obtain many results analytically in closed form.

Section III deals with various mathematical preliminaries. A brief, selective, tutorial overview of the central limit theorem, infinitely divisible distributions, and independent-increment processes is presented to place this work in perspective (as well as to fix notation). No attempt is made to be exhaustive in tbe discussion.

Tbe length of the discussion is due to tbe many special sets of parameter values that must he taken into account to be thorough. Tbe main reason for this completeness is to adequately cover all cases where uncertainty is modeled using a distribution arising from a central-limit-theorem type of argument. Tbe main contribution here is the results per se, many of which are presented bere for the first time, which unfortunately often involve eitber tedious algebraic manipulation or machine calculations. It is boped this will not obscure the surprising (at first glance) nature of the results: the quite singular bebavior of botb tbe log-likelibood-ratio test and (perhaps more importantly) its performance, for the gaussian vs nongaussian stable distribution, in distinguishing eitber location or scale. The generalization of tbese two

results to a wide class of infinitely divisible distributions (which include the family of stable distributions) is immediate, and is sketched at the end of Section IV.

III. MATHEMATICAL PRELIMINARIES

The reader is assumed to be familiar with the fundamentals of measure theory and probability theory, as found in standard references.[9–12]

Underlying the discussion to follow are:

(*i*) The notion of a probability space: a triple $\{\Omega, A, P\}$, where Ω is tbe set of elementary events, A is a σ-algebra of Borel measurable subsets of Ω, and P is a probability measure on A.

(*ii*) The definition of a stochastic process $x(t, \omega)$ defined on a parameter set E (henceforth called time), with $t \in E$, $\omega \in \Omega$, which is a function mapping the direct product $E \times \Omega$ into the real line, and the associated probability measure induced by $x(t, \omega)$.

(*iii*) Tbe measure theoretic concept of absolute continuity of one measure with respect to another, and the measure tbeoretic Lebesgue decomposition tbeorem.

3.1 *Infinitely divisible distributions and independent-increment processes*

In this section, various properties of infinitely divisible distributions and independent-increment processes are briefly reviewed. The interested reader is referred to the literature for much more information.[12–15]

This tutorial section serves several purposes:

(*i*) It gatbers together for convenient reference all material on stable distributions to be used in Part II.

(*ii*) It fixes notation.

(*iii*) It emphasizes the central role played by stable distributions in understanding both the central limit theorem and the Lévy decomposition of the infinitely divisible distributions.

(*iv*) Finally, it alerts the reader to the ricb structure and variety of infinitely divisible distributions, in general, and stable distributions, in particular, in the hope that tbey will find greater use in modeling uncertainty.

The characteristic function of a (first-order) probability distribution $P(x)$* is defined as

$$C_x(v) = \int e^{ivx} dP(x) = E(e^{ivx}) \qquad \text{a.s.}$$

* Upper case $P(\cdot)$ will denote a probability distribution, while lower case $p(\cdot)$ will denote the associated probability density function; all probability distributions examined here in any detail are absolutely continuous with respect to Lebesgue measure.

It can be shown that two probability distributions are identical if and only if their characteristic functions are identical (Ref. 14, page 28); thus, there is a one-to-one correspondence between characteristic functions and probability distribution functions. A random variable is said to be infinitely divisible if, for every natural number n, the random variable can be represented as the sum of n independent identically distributed (i.i.d.) random variables, or equivalently if its characteristic function can be written as

$$C_x(v) = [\tilde{C}_x(v, n)]^n \qquad n = 1, 2, \cdots,$$

where $\tilde{C}_x$ is the characteristic function of some probability distribution which may depend on n. Two well-known examples of infinitely divisible random variables are the gaussian [taking values on $(-\infty, \infty)$] and the Poisson (taking values at nonnegative integer multiples of h):

$$\text{Gaussian:} \quad C_x(v) = \int_{-\infty}^{\infty} e^{ixv} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\{-(x-m)^2/2\sigma^2\} dx$$
$$= \exp(imv - \tfrac{1}{2}\sigma^2 v^2)$$
$$\text{Poisson:} \quad C_x(v) = \sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} (e^{ivh})^k = \exp[\lambda(e^{ivh} - 1)].$$

De Finetti conjectured that any infinitely divisible distribution could be written as the convolution of a gaussian and a generalization of the Poisson; the resulting characteristic function can be written as

$$\ln C_x(v) = imv - \tfrac{1}{2}\sigma^2 v^2 + \int (e^{ivu} - 1) dF(u),$$

where the measure $F(u)$ specifies at what points the Poisson variable takes on nontrivial values. However, this conjecture was shown to hold only for a subset of the infinitely divisible distributions by Lévy; if one desires a canonical form of the characteristic function of an infinitely divisible distribution, then the following remarkable theorem can be proved (Ref. 13, page 76).

Theorem (Lévy): Any infinitely divisible characteristic function can be uniquely written in the canonical form

$$\ln C_x(v) = i\delta v - \tfrac{1}{2}\sigma^2 v^2 + \int_{-\infty}^{0-} \left(e^{ivu} - 1 - \frac{ivu}{1+u^2}\right) d\nu_-(u)$$
$$+ \int_{0+}^{\infty} \left(e^{ivu} - 1 - \frac{ivu}{1+u^2}\right) d\nu_+(u),$$

where δ is a location parameter $(-\infty < \delta < \infty)$, $\sigma^2 > 0$ is the variance of the gaussian component, and (ν_-, ν_+) are called the Lévy measure of the

generalized Poisson distribution. The conditions the Lévy measure must satisfy are (i) ν_- and ν_+ are nondecreasing on the intervals $(-\infty, 0)$ and $(0, \infty)$, respectively, (ii) $\nu_-(-\infty) = \nu_+(\infty) = 0$, and (iii) for every finite $\epsilon > 0$,

$$\int_{-\epsilon}^{0-} u^2 d\nu_-(u) < \infty \qquad \int_{0+}^{\epsilon} u^2 d\nu_+(u) < \infty .$$

Some examples now follow:

Example 1 (Poisson): $\delta = ivh\lambda/(1 + h^2)$, $\sigma^2 = 0$, $\nu_- = 0$, $-\infty < u < 0$

$$\nu_+ = \begin{cases} -\lambda & 0 < u < h \\ 0 & h \leqq u < \infty \end{cases};$$

$$\therefore \quad \ln C_x(v) = \lambda(e^{ivh} - 1).$$

Example 2 (Cauchy): $\sigma^2 = 0$, $\delta = 0$,

$$\nu_- = \frac{c}{\pi|u|} \qquad -\infty < u < 0$$

$$\nu_+ = \frac{-c}{\pi u} \qquad 0 < u < \infty;$$

$$\therefore \quad \ln C_x(v) = i\delta v - c|v|.$$

Example 3 (Gamma): $\sigma^2 = 0$, $\nu_- = 0$, $-\infty < u < 0$

$$\delta = p \int_0^\infty \frac{e^{-qu}}{1 + u^2}\, du < \infty$$

$$d\nu_+(u) = pe^{-qu} d(\ln u);$$

$$\therefore \quad C_x(v) = \left(1 - \frac{iv}{p}\right)^{-q}.$$

Most of the attention here will be focused on one particular class of infinitely divisible distributions, the stable distributions.

Definition: A probability distribution is said to be stable if, for all $a_1 > 0$, $a_2 > 0$, b_1, b_2, there exist constants $a > 0$, b such that

$$P(a_1 x + b_1) * P(a_2 x + b_2) = P(ax + b),$$

where $*$ denotes convolution. In other words, stable distributions are closed under the action of the group of linear affine transformations on the real line.

An important reason for examining stable distributions is found in the central limit theorem (Ref. 13, page 162; Ref. 15, page 168):

Theorem: *$P(x)$ is a limiting distribution for a sum of suitably scaled and translated, independent, identically distributed, random variables if and only if $P(x)$ is stable.*

In many practical problems, a large number of independent disturbances add and introduce uncertainty in a measurement. To analyze the effects of uncertainty, it is often convenient to replace this sum by its limiting distribution, whicb must be a stable distribution. The reader is referred to the bibliography for references on exactly what conditions govern the limiting distribution being gaussian vs nongaussian stable (Ref. 12, pages 171–190; Ref. 15, pages 165–169).

Stable distributions are infinitely divisible; the associated Lévy measures can be shown to be $\nu_-(u) = c_-|u|^{-\alpha}$, $\nu_+(u) = -c_+u^{-\alpha}$ (Ref. 13, pages 164–168; Ref. 14, pages 128–133). Requirement (*i*) tbat tbe measure be nondecreasing leads to $\alpha > 0$, while tbe final requirement (*iii*) forces $\alpha < 2$. Substituting this into the canonical representation of the characteristic function of an infinitely divisible distribution and explicitly evaluating the integral over the Lévy measure results in the following theorem:

Theorem (Ref. 13, page 164; Ref. 14, page 136): The characteristic function of a stable distribution can be expressed as

$$\ln E(e^{ixv}) = \begin{cases} -\gamma|v|^{\alpha}\left[1 + i\beta\dfrac{v}{|v|}\tan\left(\dfrac{\pi\alpha}{2}\right)\right] + i\delta v & \alpha \neq 1, \\ -\gamma|v|\left[1 + i\beta\dfrac{v}{|v|}\dfrac{2}{\pi}\ln|\gamma v|\right] + i\delta v & \alpha = 1, \end{cases}$$

where $0 < \alpha \leqq 2$, $-1 \leqq \beta \leqq 1$, $\gamma > 0$ $(\gamma \doteq c^{\alpha})$, $-\infty < \delta < \infty$. *For* $0 < \alpha < 1$, $\beta = c_- - c_+/c_- + c_+$; *for* $1 \leqq \alpha \leqq 2$, $\beta = c_+ - c_-/c_+ + c_-$. Note that for $\alpha = 2$, the characteristic function, as a complex-valued function of v, is C^{∞}, but for $1 < \alpha < 2$, it is only C^1, and for $0 < \alpha \leqq 1$ is only C^0.

For fixed β $(\beta \neq 0)$, the characteristic function is discontinuous (as a function of α) in the neighborhood of $\alpha = 1$. One approach to this problem is to rewrite tbe cbaracteristic function $(\alpha \neq 1)$ as

$$\begin{aligned} \ln E(e^{ixv}) &= -\gamma|v|^{\alpha}\left[1 + i\beta\frac{v}{|v|}\tan\left(\frac{\pi\alpha}{2}\right)\right] \\ &\qquad + iv\left(\delta + \gamma\beta\tan\frac{\pi\alpha}{2} - \gamma\beta\tan\frac{\pi\alpha}{2}\right) \\ &= -\gamma|v|^{\alpha} + i\gamma\beta v\tan\frac{\pi\alpha}{2}\left[1 - |v|^{\alpha-1}\right] \\ &\qquad + iv\left(\delta + \gamma\beta\tan\frac{\pi\alpha}{2}\right). \end{aligned}$$

If a new parameter $\delta' \doteq \delta + \gamma\beta\tan(\pi\alpha/2)$ is defined, then for β fixed

$$\lim_{\alpha\to1}\tan\frac{\pi\alpha}{2}\left[1 - |v|^{\alpha-1}\right] = \frac{2}{\pi}\ln|v|.$$

By inspection, this form of the characteristic function is not discontinuous in the neighborhood of $\alpha = 1$.

Since the characteristic function is in $L_1(-\infty, \infty)$, all stable distrihutions are absolutely continuous with respect to Lehesgue measure, and have analytic prohahility density functions. Four parameters completely specify a stahle distribution:

(*i*) α, the characteristic index of the stable distribution $P(X; \alpha, \beta)$ is associated with the asymptotic behavior of $P(X; \alpha, \beta)$. For $-1 < \beta < 1, 0 < \alpha < 2$,

$$\lim_{X \to -\infty} |X|^\alpha P(-X) = k_- > 0, \qquad \lim_{X \to \infty} X^\alpha [1 - P(X)] = k_+ > 0.$$

For $\beta = -1$ (a similar argument holds for $\beta = +1$), Lipschutz[16] and Ihragimov and Linnik (Ref. 17, pages 62 to 66)* have shown that for $1 < \alpha < 2$,

$$P(X) = 0\{k(\alpha)|X|^{\alpha/2(1-\alpha)} \exp[-c(\alpha)|X|^{\alpha/\alpha-1}]\} \quad \text{as} \quad X \to -\infty$$
$$\lim_{X \to \infty} X^\alpha [1 - P(X)] = k_+ > 0,$$

while for $0 < \alpha < 1$,

$$P(X) = 0\{k(\alpha) X^{\alpha/2(1-\alpha)} \exp[-c(\alpha) X^{-\alpha/1-\alpha}]\} \quad \text{as} \quad X \downarrow 0+$$
$$\lim_{X \to \infty} X^\alpha [1 - P(X)] = k_+ > 0,$$

where $k(\alpha)$, $c(\alpha)$ are constants which depend only on α. For the asymmetric Cauchy probability density function, it can he shown (Ref. 17, pages 57 to 60) that

$$p(X; \alpha = 1, \beta = -1) = 0\left[\exp\left(\frac{\pi}{4}|X| - \frac{2}{\pi e}\exp(\pi|X|/2)\right)\right] X \to -\infty$$
$$\lim_{X \to \infty} p(X; \alpha = 1, \beta = -1) X^2 = k_+ > 0.$$

(*ii*) β characterizes skewness of the distribution: if $\beta = 0$ the distribution is symmetric about $x = \delta$. Otherwise,

$$\lim_{X \to \infty} \frac{1 - P(X; \alpha, \beta) - P(-X; \alpha, \beta)}{1 - P(X; \alpha, \beta) + P(-X; \alpha, \beta)} = -\beta$$
$$\lim_{X \to \infty} \frac{P(-X; \alpha, \beta)}{1 - P(X; \alpha, \beta)} = \frac{1 + \beta}{1 - \beta}.$$

For $1 < \alpha < 2$, the distrihution is skewed to the left for $-1 \leqq \beta < 0$, since $P(\delta) < 1 - P(\delta)$, with the degree of skewness increasing as β

* Note typographical errors in eqs. (2.4.30) and Theorem (2.4.7), of Ref. 17.

decreases. It suffices to consider varying β over one half its range because from the characteristic function it follows that the probability density $p(x)$ obeys the relation

$$p(x;\alpha,\beta,\gamma,\delta=0)=p(-x;\alpha,-\beta,\gamma,\delta=0).$$

(iii) γ(or $c \triangleq \gamma^{1/\alpha}$) is a measure of the dispersion or spread of the distribution.

(iv) δ is a location parameter, and for $1<\alpha\leqq 2$, δ is the mean.

Only three analytic closed-form expressions for stable probability density functions are known at present:

Gaussian $(\alpha=2,\ -1\leqq\beta\leqq 1)$:

$$p(x)=\frac{1}{\sqrt{4\pi c^2}}\exp\left[-\left(\frac{x-\delta}{2c}\right)^2\right]\qquad -\infty<x<\infty;$$

Cauchy $(\alpha=1,\ \beta=0)$:

$$p(x)=\frac{c}{\pi}\left[(x-\delta)^2+c^2\right]^{-1}\qquad -\infty<x<\infty;$$

Pearson V $(\alpha=\frac{1}{2},\ \beta=-1)$:

$$p(x)=\begin{cases} c\dfrac{1}{\sqrt{2\pi}}\left(\dfrac{x-\delta}{c}\right)^{-\frac{3}{2}}\exp\left[-\dfrac{c}{2(x-\delta)}\right] & x\geqq\delta\\ 0 & x<\delta\end{cases}$$

and its conjugate density

$$p(x;\alpha=\tfrac{1}{2},\beta=1,\gamma,\delta=0)=p(-x;\alpha=\tfrac{1}{2},\beta=-1,\gamma,\delta=0).$$

Series expansions are known for the remaining stable density functions (Ref. 14, pages 138–148):

$$p(x;\alpha,\beta,\gamma=1,\delta=0)=\frac{1}{\pi}\sum_{k=1}^{\infty}\frac{(-1)^k\Gamma\left(\dfrac{k}{\alpha}+1\right)}{k!}x^{k-1}\sin\frac{k\pi}{2\alpha}(\theta-\alpha)\qquad 1<\alpha\leqq 2,$$

$$p(x;\alpha,\beta,\gamma=1,\delta=0)=\frac{1}{\pi}\sum_{k=1}^{\infty}\frac{(-1)^k\Gamma(k\alpha+1)}{k!}x^{-\alpha k-1}\sin\frac{k\pi}{2}(\theta-\alpha)\qquad 0<\alpha<1,$$

$$p(x;\alpha,\beta,\gamma=1,\delta=0)=\frac{1}{\pi}\sum_{k=0}^{\infty}\frac{(-1)^k}{k!}x^k\left[\int_0^{\infty}t^k\{\sin(1+\beta)t\}e^{-(2\beta/\pi)t\ln t}dt\right]^*\qquad \alpha=1,$$

where

$$\tan(\theta\pi/2)=\beta\tan(\pi\alpha/2),\qquad\text{and}\qquad x>0.$$

---

* For asymptotic expansions for $\alpha=1$, see Ref. 17, Theorem 2.4.3 and Ref. 18.

The reader can check that the series for $\alpha = 2$ reduces to the series for the gaussian, and the series for $0 < \alpha < 1$ and $|\beta| = 1$ are zero on a half line (cf. Pearson V). For $(0 < \alpha < 1, \ -1 < \beta < 1)$ and $(1 \leqq \alpha \leqq 2, \ -1 \leqq \beta \leqq 1)$, stable probability densities have support on $(-\infty, \infty)$. The series expansion for the density for $0 < \alpha < 1$ can be used as an asymptotic expansion for the density for $1 < \alpha < 2$ for $|\beta| \neq 1$. It can be shown from the characteristic function directly that all stable distributions are unimodal (Ref. 13, pages 158 to 161; Ref. 17, pages 66 to 76).

Figure 1 is a plot of various stable probability density functions for fixed α $(1 < \alpha < 2)$ and several β; for α near two, it is quite difficult to distinguish symmetric $(\beta = 0)$ and asymmetric stable distributions. Figure 2 shows that around the mode, all stable distributions appear roughly gaussian, for $1 < \alpha < 2$ (note the logarithmic scale).

For α in the neighborhood of two, the gaussian and nongaussian stable distributions are virtually identical around their mode, and it is only in the tails of these distributions that the differences are pronounced. One crude measure of the point at which the gaussian and nongaussian stable distributions diverge is the point at which the first term in the asymptotic series $(\alpha < 2)$ equals the gaussian density: for $\alpha = 1.90, 1.95, 1.99$, this occurs at 3.342, 3.635, 4.158 gaussian standard deviations, respectively.

One reason stable distributions have attracted little attention in the mathematical modeling of uncertainty is found in the theorem from Ref. 14, page 169: *A stable distribution with characteristic index α has all absolute moments of order p, $0 < p < \alpha < 2$: $E(|x|^p) < \infty$. Conversely, $E(|x|^p)$ does not exist, i.e., it diverges, for $p \geqq \alpha$, $\alpha < 2$.*

This suggests (albeit heuristically) that stable distributions may find application in modeling uncertainty when, as the number of observations increases, for $0 < \alpha < 1$, both the sample mean and sample variance "wander erratically," being dominated by one or a few observations, while for $1 < \alpha < 2$, the sample mean stabilizes but the sample variance does not [cf. Refs. 1, 2, 3, 4].

The generalization of these ideas from discrete time sequences of independent, identically distributed, random variables drawn from an infinitely divisible distribution to continuous time sample functions of an independent increment process is clear. The characteristic functional of a stationary independent increment process can be uniquely written as

$$\ln E[e^{iv[x(t)-x(s)]}]$$
$$= (t - s)\left[i\delta v - \tfrac{1}{2}\sigma^2 v^2 + \int_{-\infty}^{0-} \left(e^{ivu} - 1 - \frac{ivu}{1+u^2}\right) d\nu_-(u) + \int_{0+}^{\infty} \left(e^{ivu} - 1 - \frac{ivu}{1+u^2}\right) d\nu_+(u)\right]$$

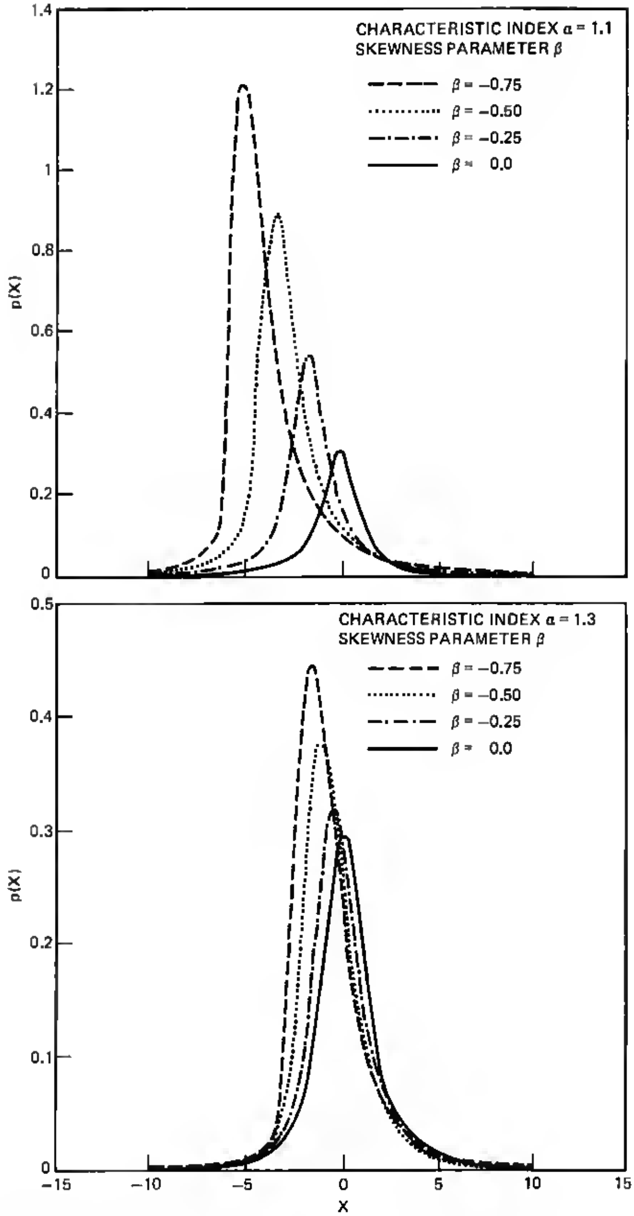

Fig. 1—Stable probability density functions [$\alpha = 1.1(0.2)1.7, \beta = -0.75(0.25)0.0$]; scale factor $c = 1.0$; location parameter $\delta = 0.0$.

for $0 \leqq s < t < T$. The parameters δ, σ^2, and (ν_-, ν_+) have been defined already. In words, any independent increment process can be decomposed into

(*i*) A singular piece, called the drift, specified by δ.

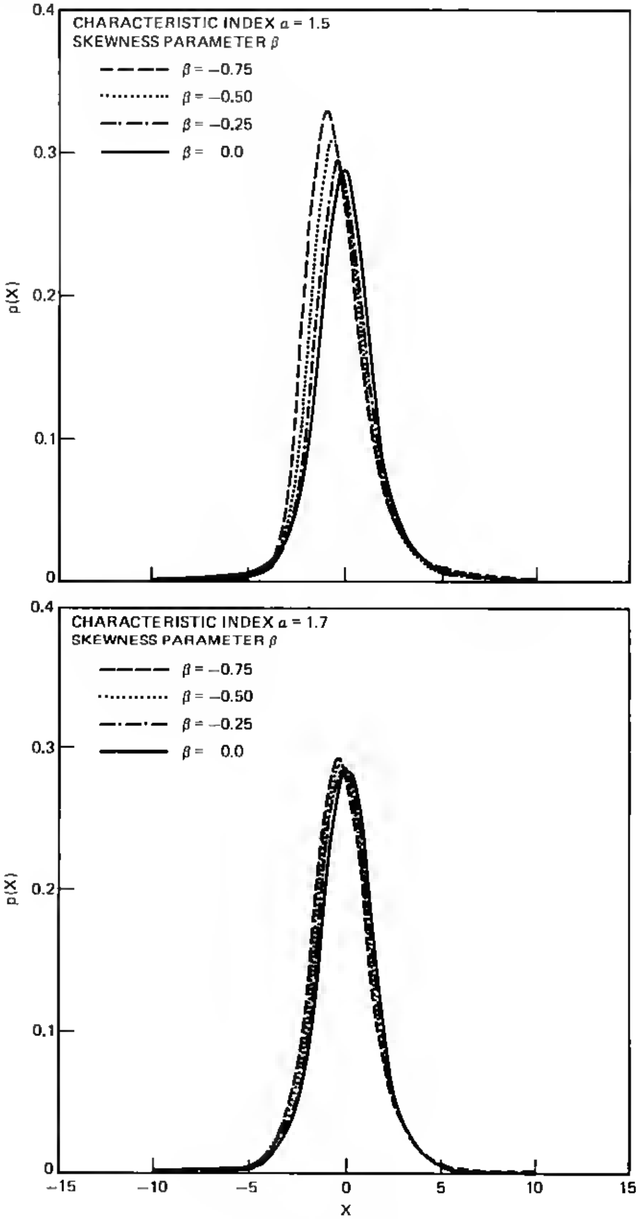

Fig. 1—(continued)

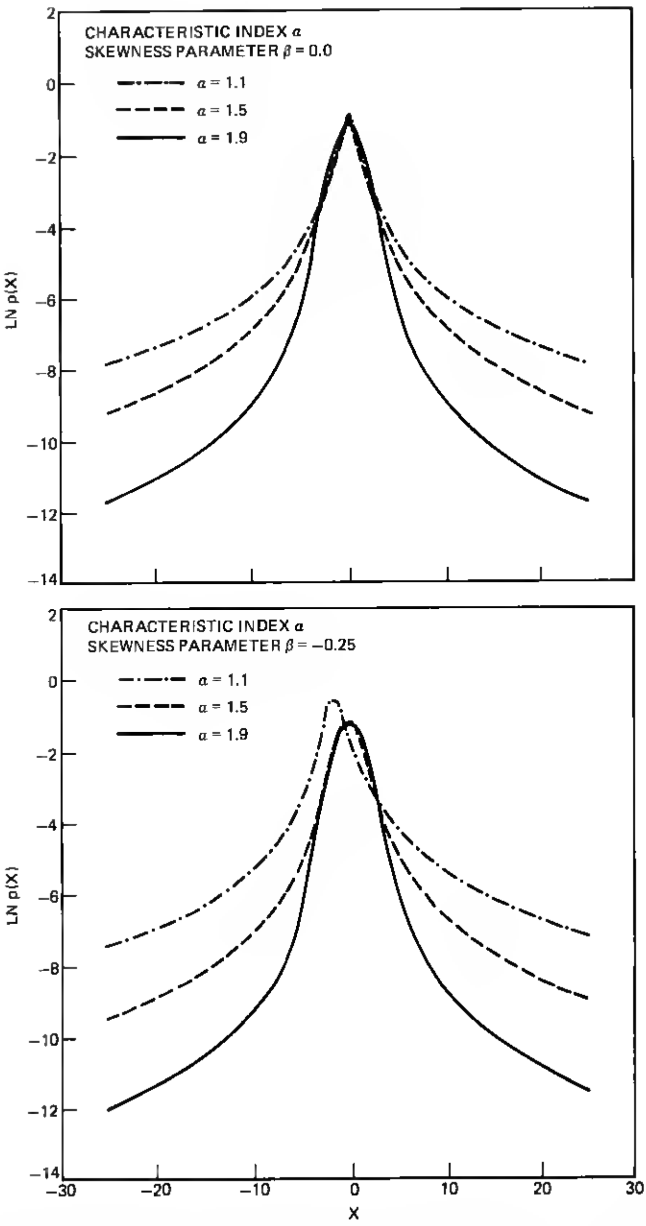

Fig. 2—Stable probability density functions (semilogarithmic) [$\alpha = 1.1(0.4)1.9$, $\beta = -0.5(0.5)0.5$]; scale factor $c = 1.0$; location parameter $\delta = 0.0$.

(*ii*) A gaussian component, a component with continuous sample paths that have unbounded variation with probability one (w.p.1), specified by σ^2.

(*iii*) A generalization of the Poisson process called a jump process,

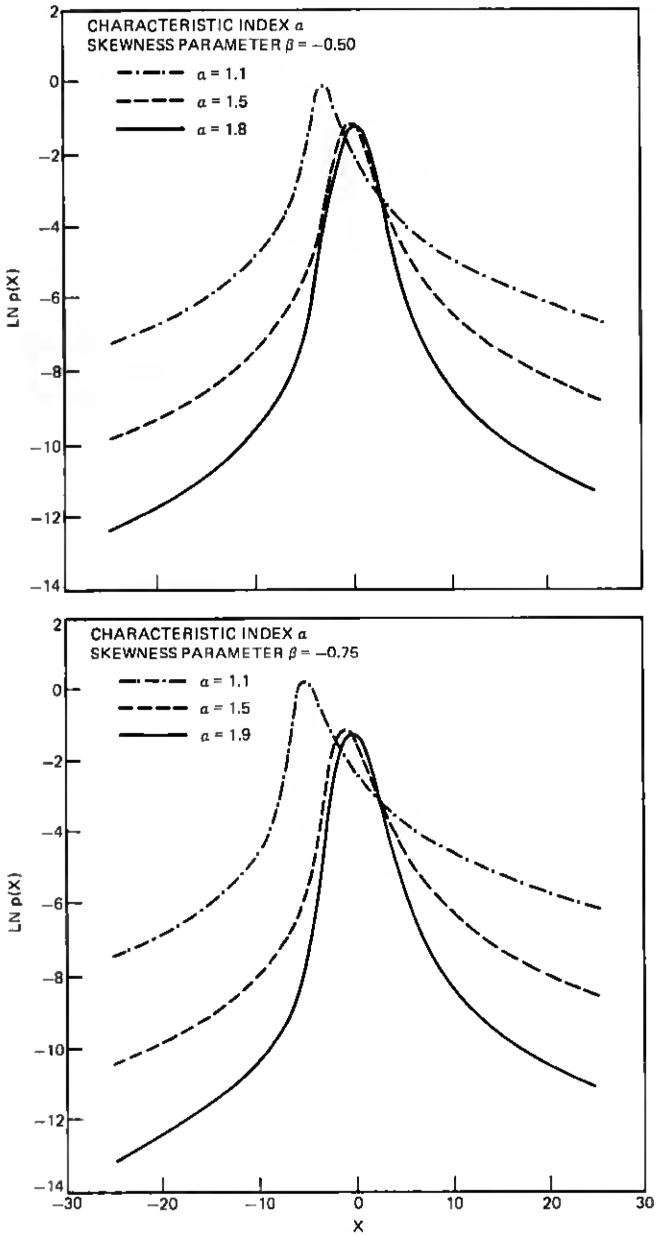

Fig. 2—(continued)

with sample paths that are constant except for simple jump discontinuities at random times with random amplitudes, specified by the Lévy measure (ν_-, ν_+).

A (separable) pure jump process, a stationary independent increment process with no gaussian component, has sample functions that are of bounded variation* with probability one if and only if

$$\int_{-1}^{0-} |u| d\nu_-(u) + \int_{0+}^{1} u d\nu_+(u) < \infty.$$

An example of an independent increment process with bounded variation (w.p.1) is a stable independent increment process $(0 < \alpha < 1)$ while stable independent increment processes $(1 \leqq \alpha \leqq 2)$ have unbounded variation (w.p.1). The intuitive meaning of the Lévy measure is that first proposed by De Finetti: the Lévy measure specifies the density of the amplitudes of the jumps of the Poisson process, provided the process sample paths are of bounded variation (w.p.1).

By allowing δ, σ^2, and (ν_-, ν_+) to depend upon time, a time-varying generalization of infinitely divisible distributions or nonstationary independent increment processes is obtained. By examining nonanticipative functionals of either a discrete time sequence of i.i.d. random variables drawn from an infinitely divisible distribution, or a continuous time independent increment process, a wide variety of Markov processes are derived. Thus, the generalizations of the results presented here to many other situations may sometimes be immediate. The richness of this class of random processes suggests these results may find wide application.

Historically, the mathematical study of independent increment processes concentrated first on the gaussian case, then on the stable case, and finally on the general case. To date, most of the engineering literature has concentrated on the gaussian case or the purely Poisson case, with the notable exception of Frost.[19] It is hoped this work will suggest promising avenues of constructive research by studying the stable case, as well as shedding light on some of the quirks of the gaussian case.

IV. DISCRETE TIME DETECTION OF TRANSLATES OF STABLE MEASURES

One of two sequences of independent, identically distributed (i.i.d.), stable, random variables is observed, under one of two hypotheses (H_0, H_1):

$$\begin{array}{lll} H_1 & r_k = s^1 + n_k & \\ & & 1 \leqq k \leqq N. \\ H_0 & r_k = s^0 + n_k & \end{array}$$

* The variation of a function $f(t)$, $0 < t < T$, is defined as $\sup \sum_{i=0}^{N-1} |f(t_{i+1}) - f(t_i)|$ where the supremum is over all possible partitions of the interval $[0, T]: 0 = t_0 < t_1 < \cdots < t_N = T$.

The observed or received sequence is denoted $\{r_k\}_1^N$, while $\{n_k\}_1^N$ is a sequence of i.i.d. stable random variables with known parameters $(\alpha, \beta, \gamma, \delta = 0)$; both s^1 and s^0 are known. The *a priori* probability of H_j is denoted π_j $(j = 0, 1)$. (The extension of allowing s^1, s^0 to depend on k is immediate and is not dealt with here.)

The measures induced by $\{r_k\}_1^N$ under H_0 and H_1 are clearly *not* mutually orthogonal. Two cases occur: for $(0 < \alpha < 1, -1 < \beta < 1)$ and $(1 \leqq \alpha \leqq 2, -1 \leqq \beta \leqq 1)$, the stable measures have support on the whole real line, and hence are equivalent. For $(0 < \alpha < 1, \beta = 1$ or $-1)$, the stable measures have support on a half line, and hence one measure is absolutely continuous with respect to the other but *not* vice versa: the supports of the two measures overlap except for the interval $[s^0, s^1)$. In either case, since the measures are not mutually orthogonal, the decision rule, which as is well known minimizes both a Bayes criterion as well as a Neyman-Pearson criterion, is the likelihood-ratio test.[5] The goal is to find the exact form of this test, and characterize its performance.* Performance here means calculating the probability that H_1 is chosen given that H_0 is true, and the probability that H_0 is chosen given that H_1 is true; these are called probabilities of error of the first and second kind, and are denoted P_{10} and P_{01}, respectively. A quantity which is also of interest is the total probability of error, defined as $(\pi_0 P_{10} + \pi_1 P_{01}) \doteq P_E$.

4.1 The likelihood-ratio test

The structure of the optimum detector is handled in two separate cases. First, when $(0 < \alpha < 1, \; -1 < \beta < 1)$ or $(1 \leqq \alpha \leqq 2, -1 \leqq \beta \leqq 1)$, the likelihood ratio is always strictly positive and finite, and is

$$\Lambda = \Lambda(r_1, \cdots, r_N) = \prod_{i=1}^{N} \frac{p_n(r_i - s^1)}{p_n(r_i - s^0)} \underset{H_0}{\overset{H_1}{\gtrless}} L,$$

where $p_n(\cdot)$ is the probability density of n_k. An equivalent test is to compute the log likelihood ratio,

$$\Lambda' = \ln \Lambda = \sum_{i=1}^{N} l(r_i) \underset{H_0}{\overset{H_1}{\gtrless}} \ln L = L',$$

where

$$l(r_i) = \ln \frac{p_n(r_i - s^1)}{p_n(r_i - s^0)},$$

and this can be explicitly calculated using the series expansions described earlier. Before doing so, it is worthwhile to examine two

* A discussion of the power of this test (or any other test) is deliberately omitted.

analytically tractable cases:

Gaussian ($\alpha = 2,\ -1 \leqq \beta \leqq 1$):

$$p_n(x) = \frac{1}{\sqrt{4\pi c}}\, e^{-x^2/4c^2} \qquad -\infty < x < \infty;$$

$$\therefore \quad l(r_i) = -\frac{1}{4c^2}\left[(r_i - s^1)^2 - (r_i - s^0)^2\right]$$

$$= \frac{1}{4c^2}\left[2r_i(s^1 - s^0) - (s^1)^2 + (s^0)^2\right];$$

$$\therefore \quad \Lambda' = \ln \Lambda = \frac{s^1 - s^0}{2c^2} \sum_{i=1}^{N} r_i - \left(\frac{N}{4c^2}\right)\left[(s^1)^2 - (s^0)^2\right] \underset{H_0}{\overset{H_1}{\gtrless}} \ln L \equiv L'.$$

The log likelihood test can be implemented using only linear processing. The rule has the interpretation of comparing an energy-like quantity, the received signal suitahly translated and squared, with a threshold. Equivalently, the test defines a hyperplane in R^N, and depending upon which side of the hyperplane $(r_1, \cdots, r_N)$ lies, H_1 or H_0 is chosen. All of this is well known (see Ref. 5, pages 94–97 and 163–173).

Cauchy ($\alpha = 1, \beta = 0$):

$$p_n(x) = \frac{c}{\pi}(x^2 + c^2)^{-1} \qquad -\infty < x < \infty;$$

$$\therefore \quad l(r_i) = \ln \frac{(r_i - s^0)^2 + c^2}{(r_i - s^1)^2 + c^2};$$

$$\therefore \quad \Lambda' = \sum_{i=1}^{N} \ln \frac{(r_i - s^0)^2 + c^2}{(r_i - s^1)^2 + c^2} \underset{H_0}{\overset{H_1}{\gtrless}} \ln L \equiv L'.$$

Unlike the gaussian case, the Cauchy log likelihood detector operates nonlinearly on the ohservation. A straightforward Taylor series expansion of the log likelihood ahout $r_i = \frac{1}{2}(s^1 + s^0)$ shows that for small perturhations ahout this point the log likelihood is linear in the perturhing quantity. On the other hand, for large excursions in any one ohservation,

$$\left|\frac{r_i - s^1}{c}\right| \gg 1, \qquad \left|\frac{r_i - s^0}{c}\right| \gg 1,$$

this one term in the sum hehaves as $O(r_i^{-1})$ or, in other words, very large excursions in the received signal are essentially (hut not entirely) discarded; this type of behavior will be called soft limiting. Only for $N = 1$ does this test reduce to finding a hyperplane and determining on which side of the hyperplane the observation lies in order to choose H_1 or H_0.

The cases ($0 < \alpha < 1$, $-1 < \beta < 1$) and ($1 \leqq \alpha \leqq 2$, $-1 < \beta < 1$), can now be examined; it is a straightforward exercise to substitute into the log likelihood the series expansions for stable probability density

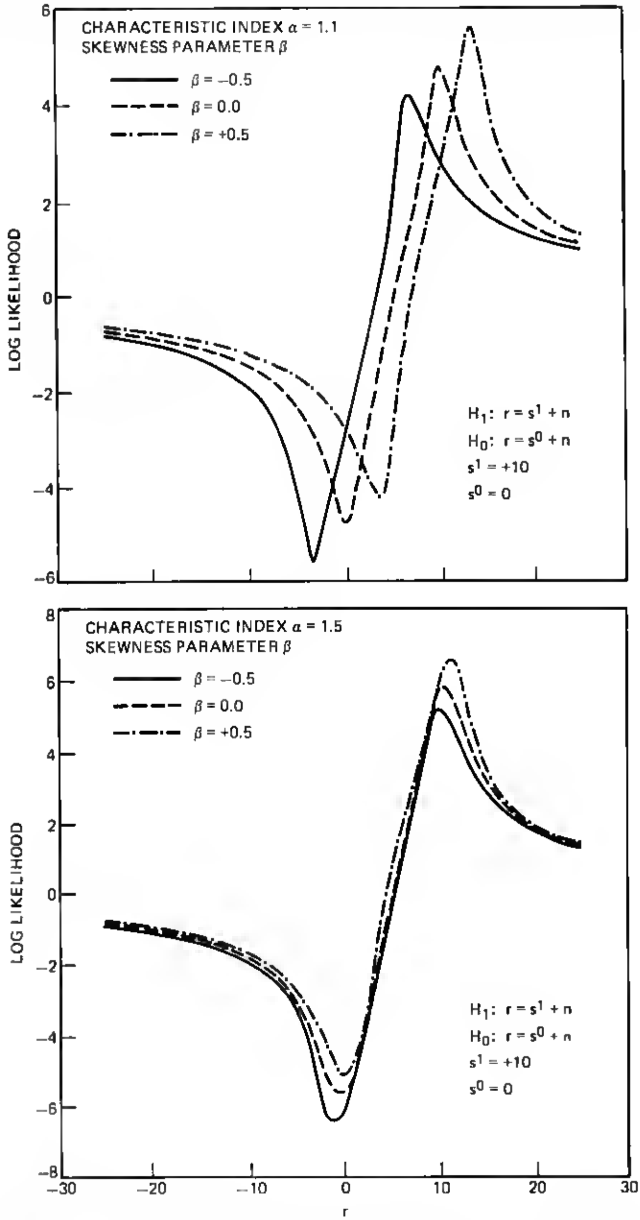

Fig. 3—Representative log likelihood functions ($s^1 = +10$, $s^0 = 0$) (α fixed, β varying); scale factor $c = 1.0$; location parameter $\delta = 0$.

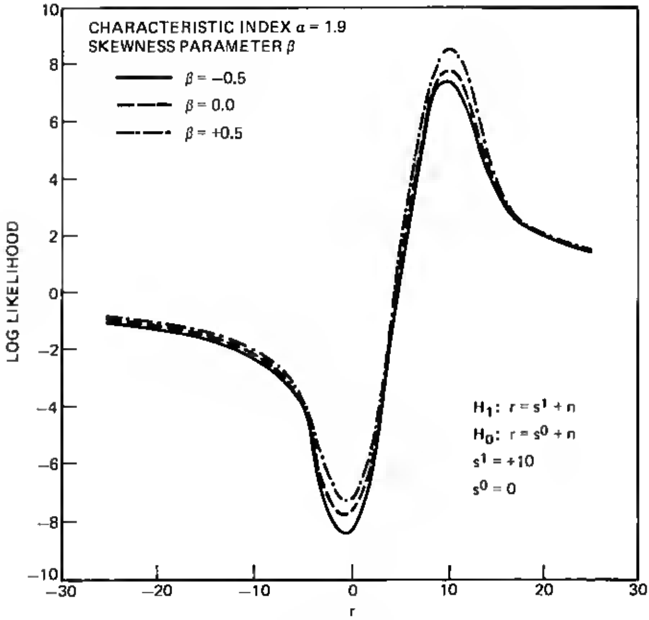

Fig. 3—(continued)

functions. Figures 3 and 4 show various representative log likelihood ratios $[l(r_i)]$ for $(1 < \alpha < 2,\ -1 < \beta < 1)$ with a fixed α and β varying; Fig. 5 shows the same log likelihood ratios as in Fig. 4 with β fixed and α varying. Similar results hold in the remaining cases $(0 < \alpha < 1,\ -1 < \beta < 1)$.

Three points are emphasized here. First, the structure of the optimum (log likelihood) detector is very sensitive to whether the underlying distribution is gaussian or nongaussian stable; this is not surprising, because small perturbations away from $\alpha = 2$ result in a singular perturbation in the probability density function.* Second, when the observation is in a neighborhood of $\frac{1}{2}(s^0 + s^1)$, an identical Taylor series argument, as used in the Cauchy example, is applicable, and small perturbations about this midway point result in linear perturbations about the corresponding log likelihood point. Third, when large excursions occur,

$$\left|\frac{r_i - s^0}{c}\right| \gg 1, \qquad \left|\frac{r_i - s^1}{c}\right| \gg 1,$$

the (log) likelihood for this term behaves as $O(r_i^{-1})$, which follows from asymptotic expansions.

* However, stable distributions in the neighborhood of $\alpha = 2$ are all close with respect to the topology induced by any reasonable metric, e.g., Prokhorov's metric.

The first two points in the preceding discussion hold for $(1 \leqq \alpha < 2, |\beta| = 1)$. The third point must be slightly modified (assume now $\beta = -1$, since a similar argument follows immediately for $\beta = 1$): $l(r_i) \sim O(r_i^{-1})$ for $r_i > 0$, but for $r_i < 0$, $l(r_i) \sim O(-|r_i|^{1/\alpha - 1})$ (cf. gaussian case) $(1 < \alpha < 2)$, while for $r_i < 0$, $\alpha = 1$,

$$l(r_i) \sim O[-\exp(\pi |r_i|/2)].$$

It remains to consider $\{n_k\}_1^N$, a sequence of i.i.d. stable random variables with $(0 < \alpha < 1, |\beta| = 1)$. Assume from here on $\beta = -1$, $s^1 > s^0$. The likelihood ratio is thus zero or strictly positive and finite, and the log likelihood is either minus infinity or finite. First, consider the Pearson V distribution as an example:

Pearson V $(\alpha = \frac{1}{2}, \beta = -1)$:

$$p_n(x) = \begin{cases} c \dfrac{1}{\sqrt{2\pi}} \left(\dfrac{x}{c}\right)^{-\frac{3}{2}} \exp[-c/2x] & x \geqq 0 \\ 0 & x < 0; \end{cases}$$

$$\therefore \quad l(r_i) = \begin{cases} -\dfrac{3}{2} \ln \left(\dfrac{r_i - s^1}{r_i - s^0} \right) - \dfrac{c}{2} \left[\dfrac{1}{r_i - s^1} - \dfrac{1}{r_i - s^0} \right] & r_i \geqq s^1 > s^0 \\ -\infty & s^1 > r_i \geqq s^0; \end{cases}$$

$$\therefore \quad \Lambda' = \sum_{i=1}^{N} \left[-\frac{3}{2} \ln \left(\frac{r_i - s^1}{r_i - s^0} \right) - \frac{c}{2} \left[\frac{1}{r_i - s^1} - \frac{1}{r_i - s^0} \right] \right] \underset{H_0}{\overset{H_1}{\gtrless}} L' \qquad r_i \geqq s^1 > s^0 \quad \text{for all } i, \quad 1 \leqq i \leqq N,$$

$$\Lambda' = -\infty \ (\text{choose } H_0) \text{ if } s^1 > r_i \geqq s^0 \quad \text{for some } i, \quad 1 \leqq i \leqq N.$$

If all the received signal samples are greater than s^1, the optimum test is to compute the log likelihood and compare it with a threshold to choose H_1 or H_0. Note that for $(r_i - s^1)/c \gg 1$, $l(r_i)$ decays asymptotically as $O(r_i^{-1})$, and thus large deviations are weighted lightly. For $r_i > s^1$, $(r_i - s^1) \ll c$, $l(r_i) \sim (r_i - s^1)^{-1}$. If one or more observations fall in the interval $[s^0, s^1)$, the optimum rule is to choose H_0.

The remaining cases $(0 < \alpha < 1$ and $\beta = -1)$ can be treated in an identical manner, using the series expansion for the densities. The important points are (*i*) the optimum detector is fundamentally nonlinear; for $(r_i - s^1)/c \gg 1$, $l(r_i)$ decays as $O(r_i^{-1})$, (*ii*) if any observation falls in the interval $[s^0, s^1)$, the optimum strategy is to choose H_0, (*iii*) for $r_i > s^1$, $|r_i - s^1| \ll c$, $l(r_i) \sim O[(r_i - s^1)^{-(1/1-\alpha)}]$.

4.2 Performance limitations

To complete the solution of the problem, the probabilities of error of the first and second kind must be calculated. This appears to be

quite difficult in the general case of an arbitrary stable distribution and bounds are developed in Section 4.3. In this section the performance of the optimum (log likelihood) detector is found explicitly for the three analytically tractahle stable distributions to illustrate the

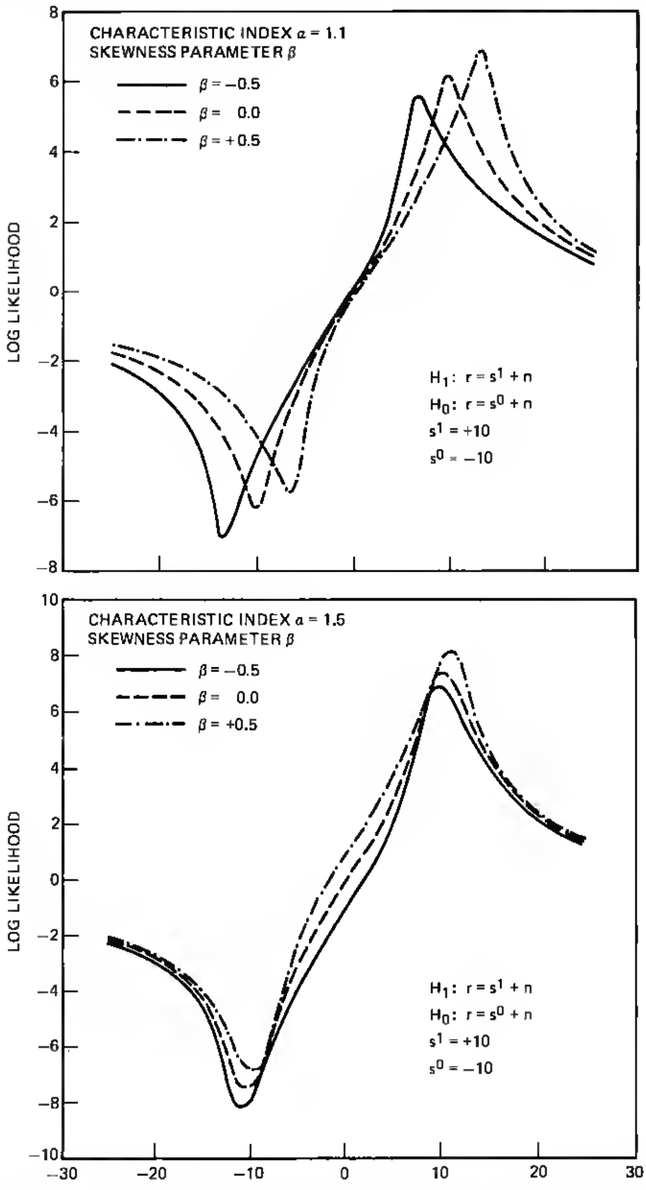

Fig. 4—Representative log likelihood functions ($s^1 = +10$, $s^0 = -10$) (α fixed, β varying); scale factor $c = 1.0$; location parameter $\delta = 0.0$.

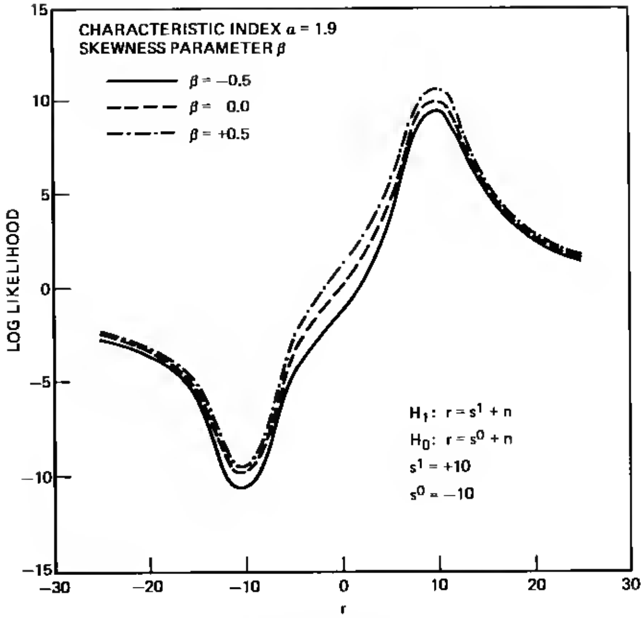

Fig. 4—(continued)

problems that must be addressed in the general case. The approach adopted is to calculate the characteristic function of the log likelihood probability measure induced under either H_1 or H_0.

Gaussian ($\alpha = 2, -1 \leqq \beta \leqq 1$)

Section 4.1 showed that the log likelihood ratio is

$$\Lambda' = -\frac{1}{4c^2}\left\{\sum_{i=1}^{N}\left[(r_i - s^1)^2 - (r_i - s^0)^2\right]\right\},$$

and since the log likelihood is a sum of i.i.d. random variables, its characteristic function can be found by using elementary Fourier techniques. The results are:

$$\ln E(e^{iv\Lambda'}|H_1) = \frac{N(s^1 - s^0)^2}{4c^2}\left[iv - v^2\right]$$

$$\ln E(e^{iv\Lambda'}|H_0) = \frac{N(s^1 - s^0)^2}{4c^2}\left[-iv - v^2\right].$$

Using the Fourier inversion lemma, the density of the log likelihood under either hypothesis can be found in closed form to be

$$p(\Lambda'|H_j) = \frac{1}{\sqrt{4\pi c'}}\exp\left[-(\Lambda' - \delta_j')^2/4c'^2\right] \qquad -\infty < \Lambda' < \infty$$

$$j = 0, 1,$$

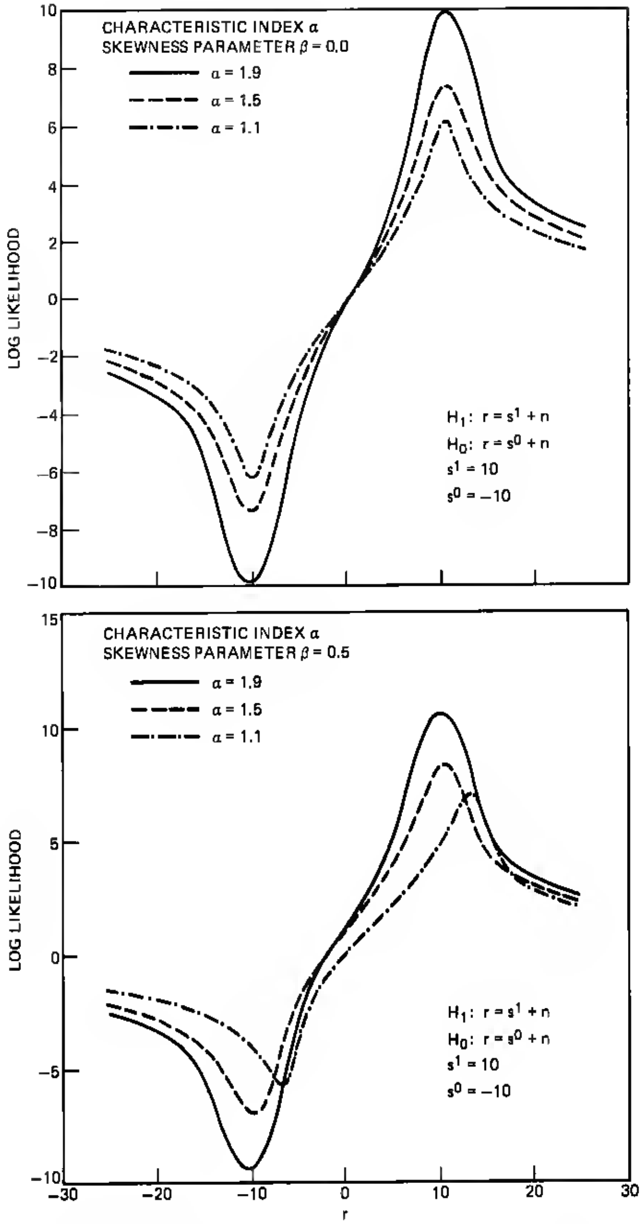

Fig. 5—Representative log likelihood functions ($s^1 = +10$, $s^0 = -10$) (α varying, β fixed).

where

$$\delta_1' = -\delta_0' = \frac{N}{4c^2}(s^1 - s^0)^2, \qquad c' = \frac{|s^1 - s^0|\sqrt{N}}{2c}.$$

The probabilities of an error of the first and second kind are

$$\begin{aligned}
P_{10} = \Pr[\text{choose } H_1 | H_0 \text{ true}] &= \int_{L'}^{\infty} p(\Lambda' | H_0) d\Lambda' \\
&= \frac{1}{2}\operatorname{erfc}\left(\frac{L' - \delta_0'}{2c'}\right) = \frac{1}{2}\left\{1 - \left(\frac{L' - \delta_0'}{c'\sqrt{\pi}}\right)\right. \\
&\quad \left.\cdot \exp\left[-\left(\frac{L' - \delta_0'}{2c'}\right)^2\right] {}_1F_1\left[1; \frac{3}{2}; \left(\frac{L' - \delta_0'}{2c'}\right)^2\right]\right\}. \\
P_{01} = \Pr[\text{choose } H_0 | H_1 \text{ true}] &= 1 - \frac{1}{2}\operatorname{erfc}\left(\frac{L' - \delta_1'}{2c'}\right) \\
&= \frac{1}{2}\left\{1 + \left(\frac{L' - \delta_1'}{c'\sqrt{\pi}}\right)\exp\left[-\left(\frac{L' - \delta_1'}{2c'}\right)^2\right]\right. \\
&\quad \left.\cdot {}_1F_1\left[1; \frac{3}{2}; \left(\frac{L' - \delta_1'}{2c'}\right)^2\right]\right\},
\end{aligned}$$

where erfc $(\cdot)$ is the complementary error function (Ref. 20, eq. 7.1.2) and ${}_1F_1$ is a hypergeometric function (Ref. 20, eq. 7.1.21; see also Slater, Ref. 21).

Cauchy ($\alpha = 1, \beta = 0$)*

It was noted previously that the log likelihood ratio can be written as

$$\Lambda' = \sum_{i=1}^{N} \ln \frac{(r_i - s^0)^2 + c^2}{(r_i - s^1)^2 + c^2}.$$

The characteristic function for the log likelihood can be found just as for the gaussian case:

$$\begin{aligned}
&E(e^{iv\Lambda'} | H_1) \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left(iv \sum_{j=1}^{N} \ln \frac{(r_j - s^0)^2 + c^2}{(r_j - s^1)^2 + c^2}\right) \prod_{j=1}^{N} \frac{c\,dr_j/\pi}{(r_j - s^1)^2 + c^2} \\
&= \prod_{j=1}^{N} \int_{-\infty}^{\infty} [(r_j - s^0)^2 + c^2]^{iv}[(r_j - s^1)^2 + c^2]^{-iv-1}\left(\frac{c}{\pi}\right) dr_j \\
&= \left\{\int_{-\infty}^{\infty} [(x + \Delta)^2 + c^2]^{iv}[(x - \Delta)^2 + c^2]^{-iv-1}\left(\frac{c}{\pi}\right) dx\right\}^N,
\end{aligned}$$

where

$$\Delta = \tfrac{1}{2}(s^1 - s^0), \qquad x = r_j - \tfrac{1}{2}(s^1 + s^0).$$

---

* The following analysis was suggested to the author by S. O. Rice; any errors in the development here are the responsibility of the author alone.

It now helps to realize

$$(x \pm \Delta)^2 + c^2 = (x^2 + \Delta^2 + c^2)\left(1 \pm \frac{2\Delta x}{x^2 + \Delta^2 + c^2}\right)$$

so that the characteristic function can he written as

$$E(e^{iv\Lambda'}|H_1) = \left[\int_{-\infty}^{\infty}\left(1 + \frac{2\Delta x}{x^2 + \Delta^2 + c^2}\right)^{iv}\left(1 - \frac{2\Delta x}{x^2 + \Delta^2 + c^2}\right)^{-iv-1} \cdot \left(\frac{c}{\pi}\right)\frac{dx}{x^2 + \Delta^2 + c^2}\right]^N$$

$$= \left[\int_{-\infty}^{\infty}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\binom{iv}{m}\binom{-iv-1}{n}\left(\frac{2\Delta x}{x^2 + \Delta^2 + c^2}\right)^{m+n} \cdot \left(\frac{c}{\pi}\right)(-1)^n\frac{dx}{x^2 + \Delta^2 + c^2}\right]^N.$$

Only even powers of $(m + n)$ contribute to the integral. This observation can be combined with the definition of the beta function (Ref. 20, eq. 6.2.1.) to show that

$$E(e^{iv\Lambda'}|H_1) = \left[\sqrt{\frac{c^2}{\Delta^2 + c^2}}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\binom{iv}{m}\binom{-iv-1}{n} \cdot \left(\frac{4\Delta^2}{\Delta^2 + c^2}\right)^{(m+n)/2}\frac{(-1)^n}{\pi}\frac{\Gamma^2\left(\frac{m+n+1}{2}\right)}{\Gamma(m+n+1)}\right]^N.$$

Substituting $(m + n) = 2l$, and using the identity (Ref. 20, eq. 6.1.18)

$$\frac{\Gamma(l + \frac{1}{2})}{\Gamma(2l)} = \frac{2\sqrt{\pi}}{4^l\Gamma(l)},$$

results in the final form of the log likelihood characteristic function assuming H_1 is true,

$$E(e^{iv\Lambda'}|H_1) = \left[-\sqrt{\frac{c^2}{c^2 + \Delta^2}}\sum_{l=0}^{\infty}\left(\frac{(\Delta/2)^2}{\Delta^2 + c^2}\right)^l \cdot \frac{(-iv)_{2l}}{(l!)^2}\,{}_2F_1(-2l, iv + 1; iv - 2l + 1; -1)\right]^N.$$

The term

$$(-iv)_{2l} = \frac{\Gamma(-iv + 2l)}{\Gamma(-iv)}$$

is standard notation for Pochhammer's symbol (Ref. 20, eq. 6.1.22). A similar expression results for the characteristic function of the log

likelihood, assuming the other hypothesis is true:

$$E(e^{iv\Lambda'}|H_0) = \left[-\sqrt{\frac{c^2}{c^2+\Delta^2}} \sum_{l=0}^{\infty} \frac{(1+iv)_{2l}}{(l!)^2} \cdot \left(\frac{(\Delta/2)^2}{\Delta^2+c^2} \right)^l {}_2F_1(-2l, iv; -iv-2l; -1) \right]^N.$$

Since tbese series converge for all v ($-\infty < v < \infty$), as well as for all (finite) values of Δ and c, the Fourier inversion lemma guarantees that a unique inverse to these transforms exists, and thus in principle the density of the log likelihood under either hypothesis is known and the probabilities of error of the first and second kind can be calculated. Numerical results are presented in a later section tbat were arrived at in exactly this manner.

Several additional observations can be made. For $N = 1$ tbe log likelihood is a random variable whose distribution has compact support on the interval

$$\ln \frac{\sqrt{\Delta^2+c^2}-\Delta}{\sqrt{\Delta^2+c^2}+\Delta} \leq \Lambda' - \left(\frac{s^1+s^0}{2} \right) \leq \ln \frac{\sqrt{\Delta^2+c^2}+\Delta}{\sqrt{\Delta^2+c^2}-\Delta}$$

and tbus the support of the log likelihood distribution for any finite number of samples, say N, is on the closed interval

$$N \ln \frac{\sqrt{\Delta^2+c^2}-\Delta}{\sqrt{\Delta^2+c^2}+\Delta} \leq \Lambda' - N\left(\frac{s^1+s^0}{2} \right) \leq N \ln \frac{\sqrt{\Delta^2+c^2}+\Delta}{\sqrt{\Delta^2+c^2}-\Delta}.$$

Since the log likelihood distribution has compact support, it is well known (Ref. 22, p. 121) that its Fourier transform bas support on the entire real axis. The second observation concerns the asymptotic ($v \gg 1$) behavior of the characteristic function of the log likelihood. Since the saddle points of the log likelihood characteristic function are at $\pm\sqrt{\Delta^2+c^2}$, stationary phase arguments[23] show that asymptotically ($v \gg 1$):

$$E(e^{iv\Lambda'}|H_1) = \left[\int_{-\infty}^{\infty} \exp\left(iv \ln \frac{(x+\Delta)^2+c^2}{(x-\Delta)^2+c^2} \right) \left(\frac{c}{\pi} \right) \frac{dx}{(x-\Delta)^2+c^2} \right]$$
$$\sim \left[\frac{c^2}{\sqrt{\pi\Delta v}} (\Delta^2+c^2)^{\frac{1}{4}} \left\{ \exp\left(iv \ln \frac{\sqrt{\Delta^2+c^2}+\Delta}{\sqrt{\Delta^2+c^2}-\Delta} + i\frac{\pi}{4} \right) + \exp\left(iv \ln \frac{\sqrt{\Delta^2+c^2}-\Delta}{\sqrt{\Delta^2+c^2}+\Delta} - i\frac{\pi}{4} \right) \right\} + O\left(\frac{i}{v} \right) \right]^N,$$

so that asymptotically the characteristic function decays as $|v|^{-N/2}$. A

similar result holds for the log likelihood characteristic function assuming H_0 is true.

An alternate approach is to calculate the Mellin transform of the likelihood probability density (for $N = 1$), then raise it to the Nth power and find the inverse transform; this was investigated without success. A direct approach, convolving the probability density of the log likelihood with itself N times, was also attempted; the resulting integrals were intractable.

Pearson V ($\alpha = \frac{1}{2}, \beta = -1$)

Assuming H_1 is true, the characteristic function of the log likelihood is

$$E(e^{iv\Lambda'}|H_1) = \int_{s^1}^{\infty} \cdots \int_{s^1}^{\infty} \exp\left\{iv \sum_{j=1}^{N}\left[-\frac{3}{2}\ln\left(\frac{r_j - s^1}{r_j - s^0}\right) - \frac{c}{2}\left(\frac{1}{r_j - s^1} - \frac{1}{r_j - s^0}\right)\right] \prod_{j=1}^{N} \frac{1}{c\sqrt{2\pi}}\left(\frac{r_j - s^1}{c}\right)^{-\frac{3}{2}} \cdot \exp\left(-\frac{c}{2(r_j - s^1)}\right)dr_j\right\}$$

$$= \prod_{j=1}^{N}\left\{\int_{s^1}^{\infty} \exp\left[-\frac{3}{2}iv\ln\left(\frac{r_j - s^1}{r_j - s^0}\right) - \frac{ivc}{2}\left(\frac{1}{r_j - s^1} - \frac{1}{r_j - s^0}\right)\right] \cdot \frac{1}{c\sqrt{2\pi}}\left(\frac{r_j - s^1}{c}\right)^{-\frac{3}{2}} \exp\left(-\frac{c}{2(r_j - c)}\right)dr_j\right\}$$

$$= \left\{\int_{\Delta}^{\infty} \exp\left[-\frac{3}{2}iv\ln\left(\frac{x - \Delta}{x + \Delta}\right) - \frac{ivc}{2}\left(\frac{1}{x - \Delta} - \frac{1}{x + \Delta}\right)\right] \cdot \frac{1}{c\sqrt{2\pi}}\left(\frac{x - \Delta}{c}\right)^{-\frac{3}{2}} \exp\left(-\frac{c}{2(x - \Delta)}\right)dx\right\}^N,$$

where $\Delta = \frac{1}{2}(s^1 - s^0)$, $x = r_j - \frac{1}{2}(s^1 + s^0)$. All attempts to simplify this expression were unsuccessful. Stationary phase arguments show that asymptotically ($v \gg 1$)

$$E(e^{iv\Lambda'}|H_1) \sim \left\{\left[\sqrt{\frac{\pi}{k_1 v}}\, k_2 \exp\left(ivk_3 + i\frac{\pi}{4}\right)\right] + O\left(\frac{1}{v}\right)\right\}^N,$$

where (k_1, k_2, k_3) are complicated functions of (c, Δ).

An attempt was made to find $E(e^{iv\Lambda'}|H_0)$, assuming no observation occurred in the interval (s^0, s^1); this approach encountered the same problems as finding $E(e^{iv\Lambda'}|H_1)$, and was unsuccessful.

It is worth noting that the log likelihood has only one maximum on the interval (s^1, ∞), for $E(e^{iv\Lambda'}|H_j)(j = 0, 1)$, and hence only one stationary point enters into the stationary phase asymptotic expression for $E(e^{iv\Lambda'}|H_j)$. It can be shown this behavior is typical of any asymmetric ($|\beta| = 1$) stable distribution. In contrast, the log likelihood has

two maxima for any stable distribution ($-1 < \beta < 1$), and hence two stationary points (cf. Cauchy).

Neither the use of Mellin transforms (instead of Fourier transforms) nor convolving the log likelihood density with itself N times made the problem any more tractable.

In the case of an arbitrary stable distribution, it appears quite difficult to find the density of the log likelihood by calculating the characteristic function of the log likelihood probability measure induced under either H_1 or H_0, because only series expansions are known at present for stable probability density functions (except for the three cases covered here). Even resorting to numerical approximation techniques poses some quite difficult problems: for $0 < \alpha < 2$, $-1 \leqq \beta \leqq 1$ (as for the Cauchy and Pearson V distributions) the log likelihood characteristic function has its support on the entire axis, and oscillates and decays asymptotically as $O[(e^{ivw_0}/\sqrt{v})^N]$ from stationary phase arguments.* To accurately approximate numerically the probabilities of error of the first and second kind from the log likelihood characteristic function, the characteristic function must be approximated and stored at a great many frequencies, and the total cost (especially due to storage) can be quite high. Furthermore, one would like to carry out calculations for many different values of (α, β, γ, δ). The storage cost plus the large number of parameter variations often desired can make this program quite expensive at present.

4.3 Analytic performance bounds

Because of analytical and numerical problems encountered in explicitly calculating the probabilities of errors of the first and second kind, as well as the total probability of error, bounds on these quantities were investigated.

Let P_1 and P_0 be probability measures defined on the same measure space (Ω, A). For $0 < q < 1$, define

$$h_q(P_1, P_0) = \left(\frac{dP_1}{d\mu}\right)^q \left(\frac{dP_0}{d\mu}\right)^{1-q} \mu,$$

where μ is any measure defined on (Ω, A) such that $\mu \gg P_1$, $\mu \gg P_0$. (An example of such a μ is $\mu = P_0 + P_1$.) This definition of h_q is seen by inspection to be independent of μ. Define

$$H_q(P_1, P_0) = \int_\Omega dh_q(P_1, P_0)$$

* Different contours of integration (e.g., path of steepest descent) were investigated without success.

as the Kakutani inner product of P_0 with P_1 (Ref. 24); the classical Hellinger integral is a special case of the Kakutani inner product, and is defined as $H_{\frac{1}{2}}(P_1, P_0)$. It is known that

$$0 \leqq H_q(P_1, P_0) \leqq 1,$$

with $H_q = 1$ iff $P_1 \equiv P_0$ a.e. The Kakutani inner product can be thought of intuitively as the amount of "colinearity" or "overlap" of two probability measures, with the larger the Kakutani inner product, the larger the "overlap." A number of useful properties of the Kakutani inner product are summarized in the following easily proven lemma[24,25]:

Lemma: *(1)* *P_0 and P_1 are mutually orthogonal (denoted $P_0 \perp P_1$), $\Leftrightarrow H_q(P_0, P_1) = 0 \Leftrightarrow h_q(P_0, P_1) = 0$*

(2) *If $0 < q < 1$, $H_q(P_0, P_1)$ is continuous in q. Four cases determine the behavior of $H_q(P_0, P_1)$ at $q = 0, 1$:*

(2a) *If P_0 and P_1 are equivalent, then $H_q(P_0, P_1)$ is continuous at $q = 0$ and $q = 1$.*

(2b) *If P_0 is absolutely continuous with respect to P_1 but not vice versa, then $H_q(P_0, P_1)$ is continuous at $q = 1$ but not at $q = 0$.*

(2c) *If P_1 is absolutely continuous with respect to P_0 but not vice versa, then $H_q(P_0, P_1)$ is continuous at $q = 0$ but not at $q = 1$.*

(2d) *If P_0 and P_1 are neither mutually orthogonal nor equivalent, then $H_q(P_0, P_1)$ is discontinuous at $q = 0$, $q = 1$.*

(3) *$H_q(P_0, P_1)$ and its logarithm are convex functions, $0 < q < 1$. The convexity is strict iff $(dP_1/dP_0)(x)$ is not constant for all $x \in \operatorname{supp}(P_0) \cap \operatorname{supp}(P_1)$.*

It is instructive to rewrite $H_q(P_0, P_1)$ in two different ways to explicitly show the relationship between the log likelihood functional and the Kakutani inner product:

$$\begin{aligned} (i)\quad H_q(P_0, P_1) &= \int \exp\{q \ln (dP_1/dP_0)\} dP_0 \\ &= E\{\exp [q \ln (dP_1/dP_0)] | H_0\}, \\ (ii)\quad H_q(P_0, P_1) &= \int \exp\{(q-1) \ln (dP_1/dP_0)\} dP_1 \\ &= E\{\exp [(q-1) \ln (dP_1/dP_0)] | H_1\}. \end{aligned}$$

(i) and *(ii)* are the Laplace transforms of the log likelihood probability density (also called the moment generating function of Λ), evaluated at q and $(q - 1)$, and assuming H_0 and H_1 are true, respectively. It is

196–197). Using Hölder's inequality, it is straightforward to show that the logarithm of H_q and, hence, H_q itself, are convex functions of q, $0 < q < 1$.

Chernoff[26] was apparently first to use $H_q(P_0, P_1)$ (where $P_0 \ll \mu$, $P_1 \ll \mu$, $\mu \doteq$ Lebesgue measure) to upper bound the probabilities of error of the first and second kind, and his work has found widespread application in the engineering and statistical literature (see also, Ref. 14, pages 517–520 and the references therein).

In the notation used here, Chernoff showed

$$P_{01} \leqq \inf_{0<q} H_q(P_0, P_1)e^{-qL'}$$

$$P_{10} \leqq \inf_{q<1} H_q(P_0, P_1)e^{-(q-1)L'},$$

where L' is the threshold in the log likelihood ratio test.

Chernoff's original ideas have been generalized in several directions. Kraft[27] obtained upper and lower bounds on the total probability of error. For some choice of L' (see also Ref. 28):

$$\tfrac{1}{2} \min (\pi_0, \pi_1) H_{\frac{1}{2}}^2(P_0, P_1) \leqq P_E \leqq (\pi_0\pi_1)^{\frac{1}{2}} H_{\frac{1}{2}}(P_0, P_1).$$

Hellman and Raviv[29] have also worked on this problem. Shannon, Gallager, and Berlekamp[25] obtained lower bounds on the probabilities of error of the first and second kind in terms of the logarithm of $H_q(P_0, P_1)$, and the first and second derivatives of the logarithm.

Here the Kakutani inner product plays two key roles, providing a check on whether or not singular or perfect detection is possible [iff $H_q(P_0, P_1) \equiv 0$], as well as giving exponentially sharp bounds on the performance of the log likelihood ratio test if detection is not singular. Since the Kakutani inner product need only be calculated at a small number of values of q to accurately numerically approximate upper and lower bounds on error probabilities, unlike calculating the probabilities of error of the first and second kind from the log likelihood characteristic function, this approach may be useful as a practical design tool because it is relatively inexpensive.

The following observations are strightforward exercises:

(*i*) When a sequence of N i.i.d. random variables is observed, $H_q(P_0, P_1) = e^{-AN}$, where A is independent of N, depending solely on P_0, P_1, and q.

(*ii*) When P_0 and P_1 are absolutely continuous with respect to Lebesgue measure, and the corresponding densities are unimodal translates of one another, then for fixed q, the larger the separation the smaller the inner product $H_q(P_0, P_1)$.

The Kakutani inner product $H_q(P_0, P_1)$ can be explicitly calculated for tbe three analytic cases discussed earlier:

Gaussian ($\alpha = 2$, $-1 \leqq \beta \leqq 1$):

$$p_n(x) = 1/\sqrt{4\pi c}\exp(-x^2/4c^2) \qquad -\infty < x < \infty$$

$$H_q(P_0, P_1) = e^{N\mu(q)},\ \mu(q) \doteq \ln \int_{-\infty}^{\infty} p_n^q(x - s^1)p_n^{1-q}(x - s^0)dx$$

$$\mu(q) = -q(1 - q)(s^1 - s^0)^2/4c^2;$$

$$\therefore \quad \mu\left(\frac{1}{2}\right) = -\left(\frac{s^1 - s^0}{4c}\right)^2.$$

Cauchy ($\alpha = 1$, $\beta = 0$):

$$p_n(x) = \frac{c}{\pi}(x^2 + c^2)^{-1} \qquad -\infty < x < \infty$$

$$H_q(P_0, P_1) = \left[\int_{-\infty}^{\infty} p_n^q(x - s^1)p_n^{1-q}(x - s^0)dx\right]^N$$

$$\int_{-\infty}^{\infty} p_n^q(x - s^1)p_n^{1-q}(x - s^0)dx = \sum_{j=0}^{\infty} - \sqrt{\frac{c^2}{c^2 + \Delta^2}}\left[\frac{(\Delta/2)^2}{\Delta^2 + c^2}\right]^j \cdot \frac{(1 - q)_{2j}}{(j!)^2}\, {}_2F_1(q, -2j; q - 2j; -1),$$

where $\Delta = (s^1 - s^0)/2$.

From tables (Ref. 30, 263.00) for elliptic integrals:

$$H_{\frac{1}{2}}(P_0, P_1) = \left\{\frac{1}{\pi}\left[\left(\frac{s^1 - s^0}{2c}\right)^2 + 1\right]^{-\frac{1}{2}} cn^{-1}\left[-1, \left[\left(\frac{s^1 - s^0}{2c}\right)^2 + 1\right]^{-\frac{1}{2}}\right]\right\}^N,$$

where $cn^{-1}(\cdot, \cdot)$ is an inverse Jacobian elliptic function.

Pearson V ($\alpha = \frac{1}{2}$, $\beta = -1$):

$$p_n(x) = \begin{cases} \dfrac{1}{c\sqrt{2\pi}}\left(\dfrac{x}{c}\right)^{-\frac{3}{2}} e^{-c/2x} & x \geqq 0 \\ 0 & x < 0 \end{cases}$$

$$H_q(P_0, P_1) = \left[\int_{-\infty}^{\infty} p_n^q(x - s^1)p_n^{1-q}(x - s^0)dx\right]^N.$$

The integral could not be expressed in any other analytic form. Since P_1 is absolutely continuous with respect to P_0, but not vice versa, $H_q(P_0, P_1)$ is continuous for $q \in (0, 1]$, and is discontinuous at $q = 0$. Apparently only in the gaussian case does the Kakutani inner product or the Hellinger integral reduce to a simple form, and for general stable distributions the problem appears to be analytically intractable at present. Thus, it seemed worthwhile to investigate numerical methods

for approximating the desired integrals. Again it seems important to emphasize that an accurate approximation of the log likelihood probability density Laplace transform under H_1 or H_0 is needed at only a small number of choices of q, so the calculations can be quite inexpensive. In the previous section, the log likelihood characteristic function had to be approximated at a great many frequencies, and the resulting computation effort and storage made that program relatively more expensive.

4.4 Numerical approximation of performance bounds

At present, three approaches have been investigated for calculating stable probability density functions. The first involves summing power series and asymptotic series,[31] the second involves quadrature of an integral representation of the density,[32] and the third uses a discrete fast Fourier transform of the characteristic function (Ref. 33, pages 35–42; and Ref. 34).

The approach used here was a combination of the first and third methods. The stable probability density function was approximated over its central region via a discrete fast Fourier transform, while asymptotic expansions were used outside this region. This approach avoids the difficulty of knowing how to merge the power series and asymptotic series (see Ref. 31).

The Kakutani inner product was broken into two integrals. The first integral was approximated by a fixed step size Romberg integration

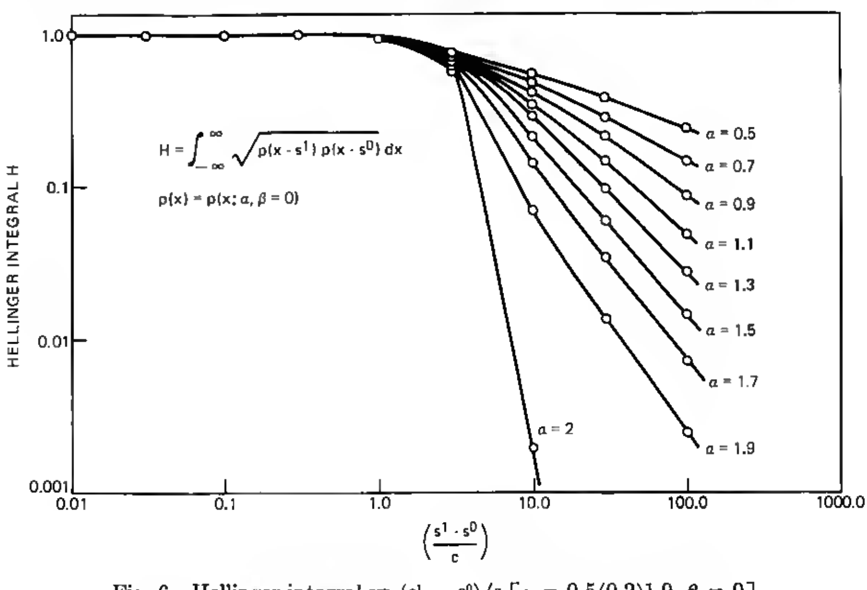

Fig. 6—Hellinger integral vs $(s^1 - s^0)/c$ $[\alpha = 0.5(0.2)1.9, \beta = 0]$.

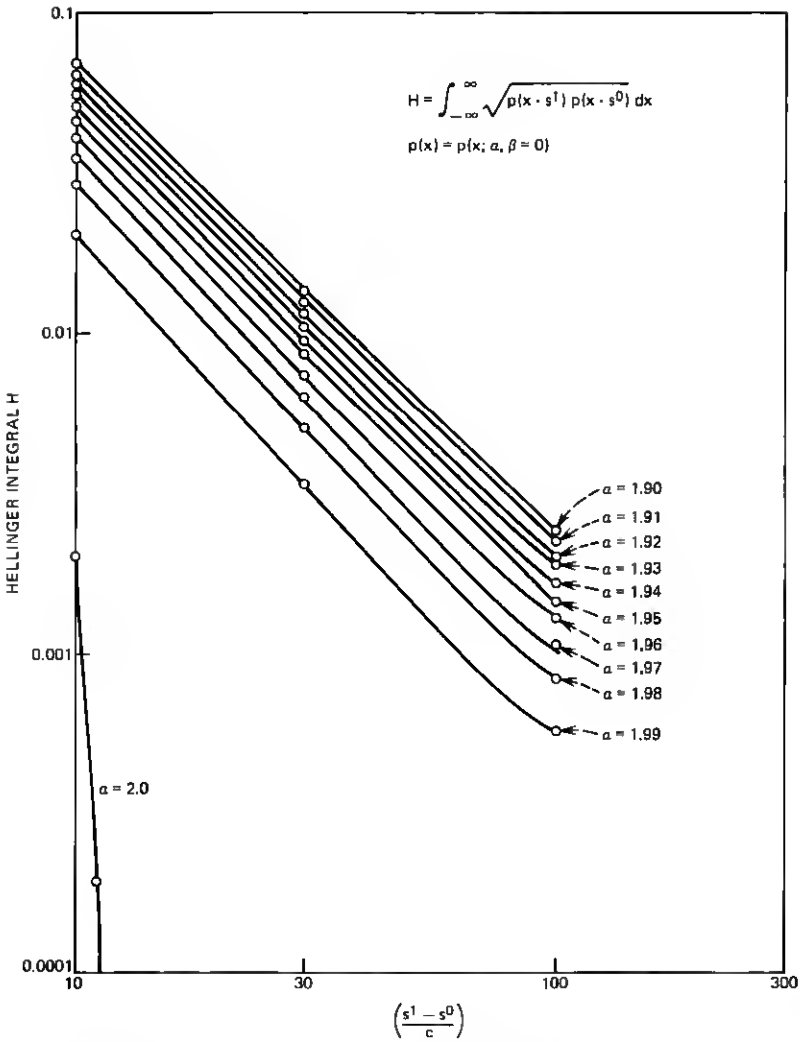

Fig. 7—Hellinger integral vs $(s^1 - s^0)/c$ [$\alpha = 1.90(0.01)2.00$, $\beta = 0$].

routine[35] using the discrete fast Fourier transform approximation to the density (typically, 4096 points were used). The second integral was approximated by a variable step size Romberg integration algorithm using the asymptotic expansion for the density.

While this approach is adequate for finite mean stable distributions ($1 < \alpha \leqq 2$), and with care works for $0.5 \leqq \alpha \leqq 1$, it is inadequate for $0 < \alpha < 0.5$, because the expense is too great at present. The reason is that for $0 < \alpha < 1$, a great many evenly spaced points must be used

to adequately approximate the characteristic function in the neighborhood of the origin (where its derivative is unbounded), as well as at otber frequencies, and the expense of storing tbese values (to carry out the discrete fast Fourier transform) is prohibitive. One possible approach around this prohlem is to simply use only the series expansion (see Ref. 33).

All results presented here were calculated on a Honeywell 6070 computer using double-precision arithmetic (14 significant figures); the estimated relative error in all cases was less than a tenth of one percent.

Figure 6 shows the Hellinger integral for various parameters $[\alpha = 0.5(0.2)1.9,\ \beta = 0]$ as a function of $[(s^1 - s^0)/c]$, for $N = 1$. This figure suggests an interesting conjecture, tbat the Hellinger integral is smaller the closer the characteristic index α is to two, all other factors heing the same. No proof of this is known, at present.

Figure 7 depicts results of numerically calculating the Hellinger integral for various characteristic indices close to two $[\alpha = 1.90(0.01)1.99,\ \beta = 0]$, for $N = 1$. The singular nature of the gaussian distribution ($\alpha = 2$) is quite evident when compared with that of $\alpha = 1.99$ or $\alpha = 1.98$.

Figure 8 shows $\mu(q)$ vs q for fixed $[(s^1 - s^0)/c]$. Again, the closer the index is to two, the smaller the inner product.

Figure 9 presents $\mu(q)$ vs q for various choices of $[(s^1 - s^0)/c]$, and fixed characteristic index α and skewness parameter β; the larger $(s^1 - s^0)/c$, the smaller $H_q(P_0, P_1)$.

4.5 Comparison of the performance of the log likelihood decision rule ($\alpha = 1.95$) with a linear decision rule

It is interesting to compare the performance of the log likelihood decision rule with a linear decision rule, when the observations are drawn from a nongaussian stahle distrihution with characteristic index near two. To be explicit, it is assumed the observations are i.i.d. stahle random variables ($\alpha = 1.95$, $\beta = 0$), with $\pi_0 = \pi_1 = \frac{1}{2}$ and $s^1 = -s^0 = S$ chosen for simplicity. The linear decision rule is simply

$$\sum_{i=1}^{N} r_i \underset{H_0}{\overset{H_1}{\gtrless}} 0.$$

This sum is a stahle random variable, with parameters ($\alpha = 1.95$ $\beta = 0$, $N\gamma$, Ns^j), assuming $H_j (j = 0, 1)$ is true. The total probability of error is equal to the prohahility of either an error of the first or second kind,

$$P_E = P_{10} = P_{01},$$

and can he computed from the series described earlier, or from pub-

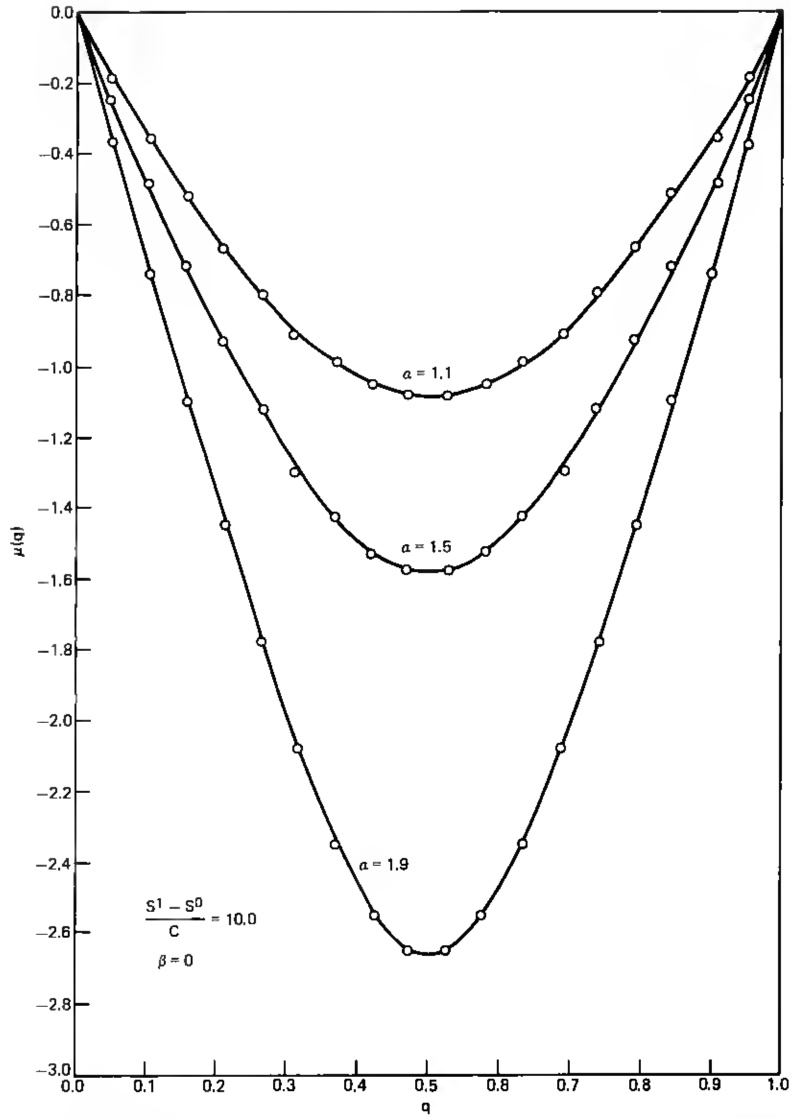

Fig. 8—Logarithm of Kakutani inner product H_q vs q [$\alpha = 1.1(0.4)1.9$, $\beta = 0$] [$(s^1 - s^0)/c = 10$].

lished tables.[31] This is plotted in Fig. 10 as a function of $[(s^1 - s^0)/c]$ for various N. The same figure includes plots of the Hellinger integral upper bound on the total probability of error using the log likelihood decision rule. The figure makes it quite clear that the log likelihood decision rule, for many cases of interest, has a much much smaller probability of error tban the linear decision rule.

Asymptotically, the total probability of error for the linear detection strategy behaves as

$$P_E \sim \mathrm{O}\left(\left[\frac{NS}{(N\gamma)^{1/\alpha}}\right]^{-\alpha}\right), \qquad \gamma \doteq c^\alpha;$$

$$\therefore \quad P_E \sim \mathrm{O}[(S/c)^{-\alpha}N^{1-\alpha}],$$

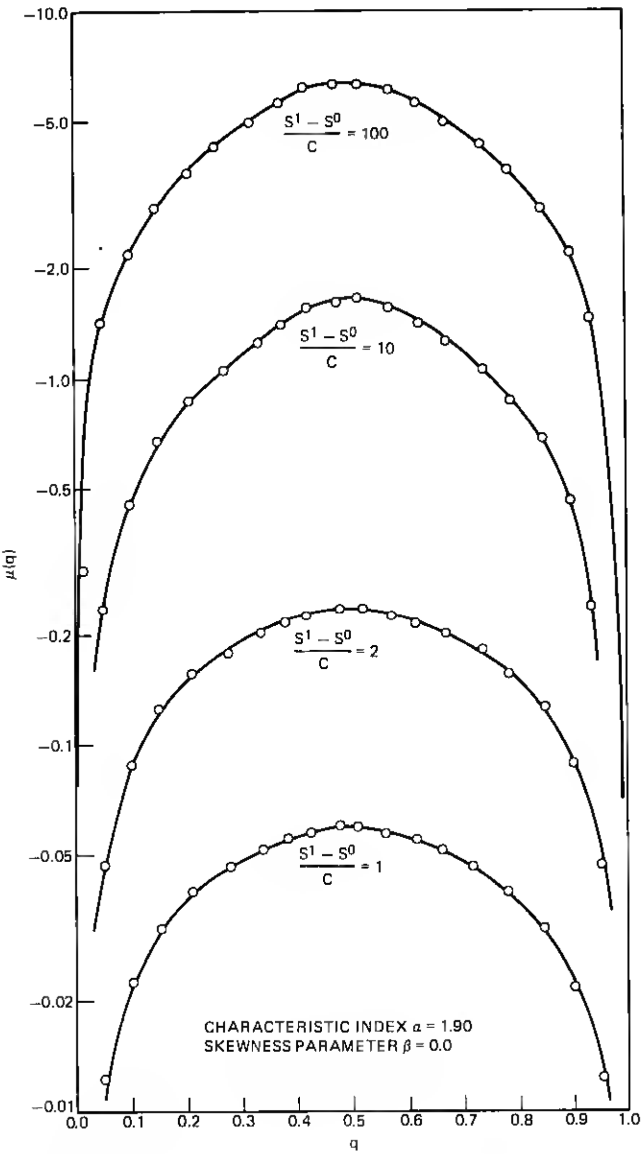

Fig. 9—Logarithm of Kakutani inner product H_q vs q $[(s^1 - s^0)/c = 1, 2, 10, 100]$ $(\alpha = 1.90, \beta = 0)$.

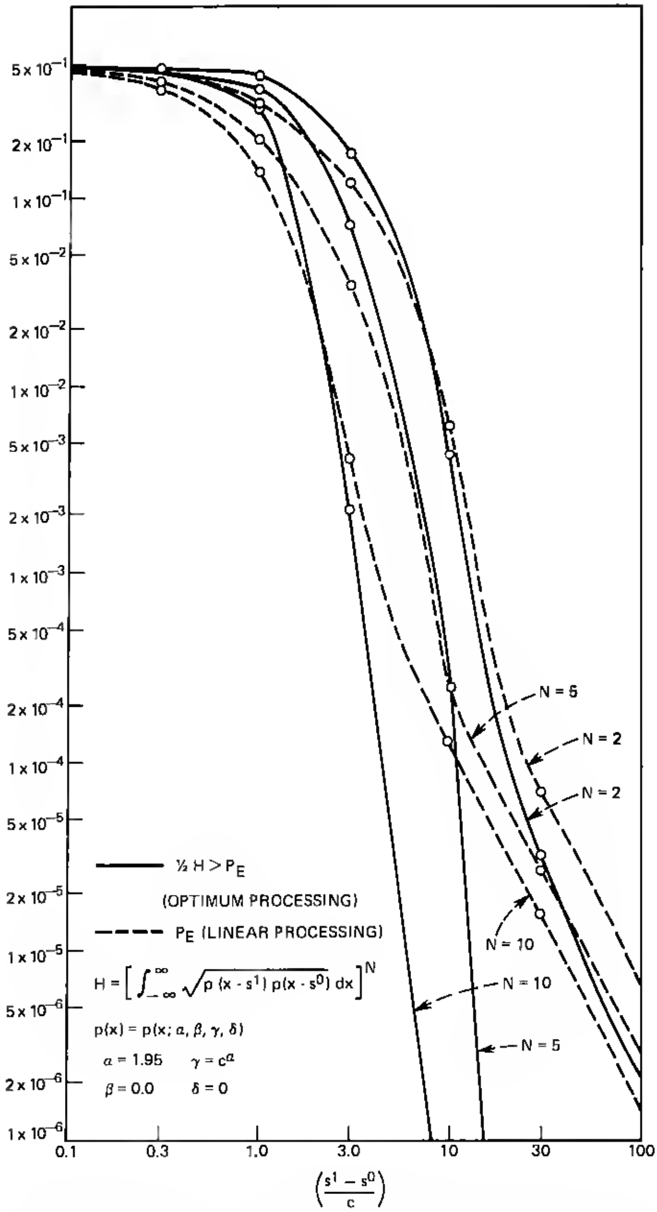

Fig. 10—Linear processing probability of error and Hellinger integral upper bound on nonlinear processing probability of error vs $(s^1 - s^0)/c$ ($\alpha = 1.95$, $\beta = 0$).

while the probability of error for the log likelihood detection strategy asymptotically behaves at

$$P_E = O(e^{-AN}),$$

where $A = A(\alpha, \beta, \gamma, S) > 0$, independent of N. This simple asymptotic analysis suggests that the log likelihood decision rule has a much smaller probability of error than the linear decision rule, for large N, which is horne out in Fig. 10.

4.6 Comparison of the upper and lower bounds and P_E

It remains to compare the hounds on total prohahility of error, and prohahilities of errors of the first and second kind, with the actual quantities. None of the hounds employed here are tight, hecause the upper and lower hounds have different exponents. This program is quite difficult, and has only been carried out analytically for the gaussian case, and numerically for the Cauchy case. The remaining cases can be handled numerically following Shannon et al.[25] For simplicity, from this point on it is assumed that $\pi_0 = \pi_1 = \frac{1}{2}$, $s^1 = -s^0 = s$.

Gaussian ($\alpha = 2$, $-1 \leqq \beta \leqq 1$)

Earlier it was shown that

$$P_E = P_{10} = P_{01} = \frac{1}{2}\,\mathrm{erfc}\left(\frac{\sqrt{N}s}{2c}\right).$$

This can be upper and lower hounded tightly hy (see Ref. 20, eq. 7.1.13)

$$K_L e^{-Ns^2/4c^2} < P_E < K_u e^{-Ns^2/4c^2},$$

where

$$K_L = \frac{1}{2}\cdot\left(\frac{\sqrt{N}s}{c} + \sqrt{\frac{Ns^2}{c^2} + 8}\right)^{-1}\cdot\pi^{-\frac{1}{2}}$$

$$K_u = \frac{1}{2}\cdot\left(\frac{\sqrt{N}s}{c} + \sqrt{\frac{Ns^2}{c^2} + \frac{16}{\pi}}\right)^{-1}\cdot\pi^{-\frac{1}{2}}.$$

Since both K_L and K_u behave as $O(N^{-\frac{1}{2}})$, $P_E \sim e^{-Ns^2/4c^2 - O[LN(N)]}$, where K_u and K_L introduce factors of log (N) in the exponent. The Hellinger integral hounds are [27]

$$\frac{1}{4}\exp\left(-N\frac{s^2}{2c^2}\right) \quad < P_E < \frac{1}{2}\exp\left(-\frac{Ns^2}{4c^2}\right).$$

By inspection, the exponent in the upper hound agrees with the tight lower and upper hound exponent [to within a factor of $LN(N)$]. The Chernoff upper hounds[26] on P_{10}, P_{01} are

$$\begin{aligned} &P_{01} \leqq \exp[-Nq^2(s/c)^2] \\ \text{or}\quad &P_{10} \leqq \exp[-N(1-q)^2(s/c)^2], \end{aligned} \qquad \text{for some } q \in [0, 1]$$

and for $q = \frac{1}{2}$ these exponents agree with the tight upper and lower

hound exponents to within a factor on $LN(N)$. The lower hounds[25] are

$$P_{01} > \tfrac{1}{4}\exp[-Nq^2(s/c)^2 - q(s/c)\sqrt{2N}]$$

$$\text{for some } q \in [0, 1]$$

$$\text{or}\quad P_{10} > \tfrac{1}{4}\exp[-N(1-q)^2(s/c)^2 - (1-q)(s/c)\sqrt{2N}],$$

and for N sufficiently large, the upper and lower hound exponents are identical within a factor of $O(N^{-\frac{1}{2}})$.

Cauchy ($\alpha = 1, \beta = 0$)

The real and imaginary parts of the characteristic function of the Cauchy log likelihood were calculated numerically at 513 evenly spaced frequencies starting at $v = 0$ from a direct numerical quadrature of the (complex) integral

$$\varphi(v) = \int_{-\infty}^{\infty} \exp\left(iv \ln \frac{(x+s)^2 + c^2}{(x-s)^2 + c^2}\right)\left(\frac{c}{\pi}\right)\frac{dx}{(x-s)^2 + c^2},$$

$$v = k\Delta v, \qquad k = 0, \cdots, 512$$

using an adaptive, step-size, Romherg, numerical integration algorithm, with an estimated error of 10^{-10} (all arithmetic was performed in douhle precision). One representative characteristic function is plotted in Fig. 11. The stationary-phase asymptotic expression was used for frequencies outside of this range. The resulting approximation to the characteristic function was multiplied hy itself N times, and a numerical approximation of the inverse transform of this resulting characteristic function was calculated, using a fixed, step-size, Romherg algorithm for the first 513 frequencies; an adaptive, step-size, Romberg algorithm was used for the tail of the inverse transform. The final results are felt to be accurate to three significant figures. The results are plotted in Fig. 12, along with the Hellinger integral upper hound. Clearly, the Hellinger integral upper hound is quite conservative; it is straightforward to check that the Hellinger integral (squared) lower hound is too optimistic, from the curves in Fig. 12.

4.7 Generalizations

The extensions of the results in this section (as well as the following section) to a much wider class of infinitely divisihle distrihutions is immediate. Here these extensions are sketched. Elementary arguments (Ref. 15, page 540) show that if the Lévy measure of an infinitely divisihle distrihution hehaves asymptotically as a power, i.e., $\nu(X, \infty) \sim O(X^{-p})$, $\nu(-\infty, -X) \sim O(X^{-q})$, then $\Pr[x > X] \sim O(X^{-p})$, $\Pr[x < -X] \sim 0(X^{-q})$, where $p, q > 0$. Given a sequence of i.i.d. random variahles drawn from such a distribution with one of two location parameters, it is straightforward to check that results analo-

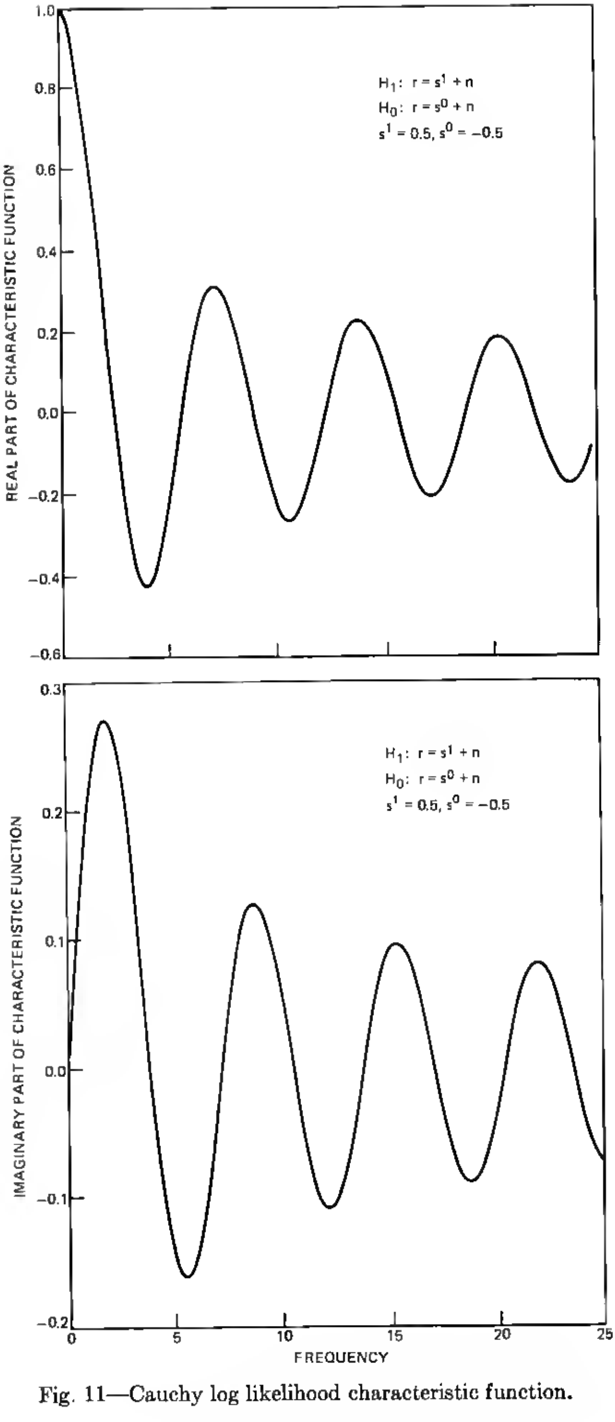

Fig. 11—Cauchy log likelihood characteristic function.

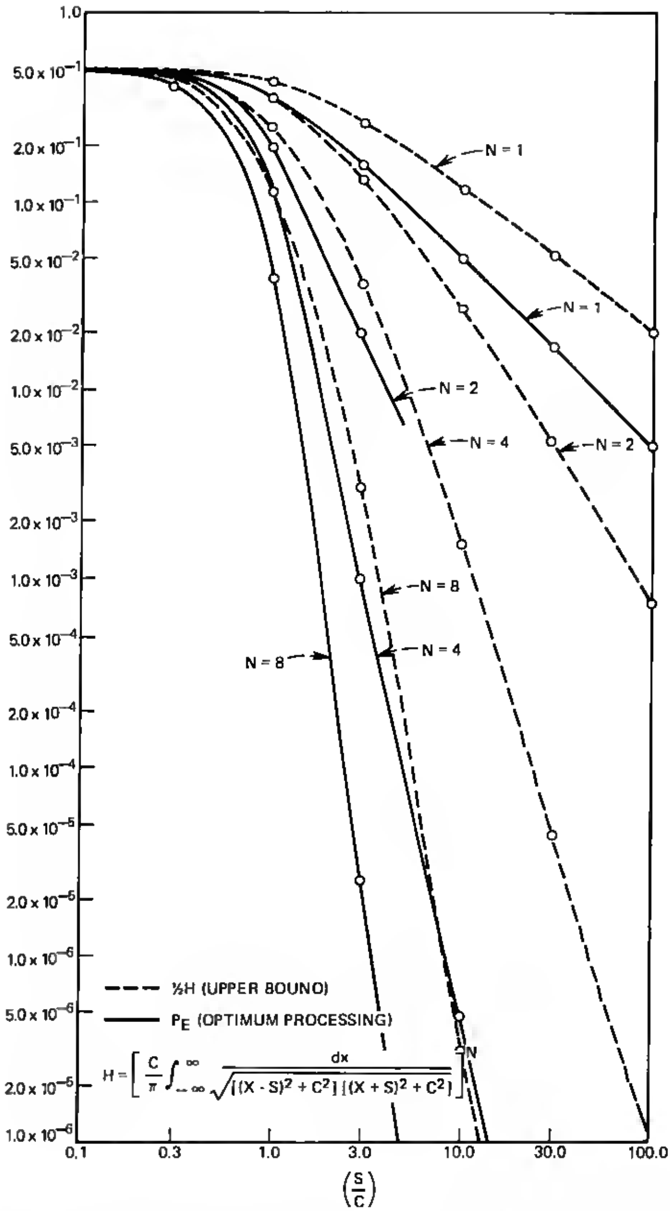

Fig. 12—Log likelihood probability of error and Hellinger integral upper bound for Cauchy ($\alpha = 1$, $\beta = 0$) samples vs (s/c).

gous to those in this section hold: (*i*) $l(r_i) \sim 0(r_i^{-1})$, (*ii*) the probability of an error of the first and second kind, using a log likelihood ratio test, is upper bounded by $\exp(-AN)$, (*iii*) using a simple linear test to discriminate between hypotheses, i.e., adding up the observations and comparing the sum with a threshold, results in the probability of an error of the first or second kind behaving as $0(NL'^{-p})$, $0(NL'^{-q})$, and choosing L' directly proportional to N (as in the gaussian case) gives P_{01}, $P_{10} \sim 0(N^{1-p})$, $0(N^{1-q})$, which is much worse than the performance of the log likelihood test in this asymptotic sense.

V. DISCRETE TIME DETECTION OF STABLE MEASURES WITH DIFFERENT SCALES

In this section, one approach is studied for hypothesis testing of different scale parameters; since the ideas are quite similar to that just developed, the treatment is much shorter.

One of two sequences of i.i.d. stable random variables is observed (under one of two hypotheses, H_0 and H_1):

$$\begin{matrix} H_1 & r_k = s^1 n_k \\ H_0 & r_k = s^0 n_k \end{matrix} \qquad 1 \leqq k \leqq N.$$

The observed or received sequence is denoted $\{n_k\}_1^N$, where the $\{n_k\}_1^N$ are i.i.d. stable random variables with known parameters (α, β, $\gamma = 1$, $\delta = 0$); both s^1 and s^0 are known. The *a priori* probability of H_j is π_j ($j = 0, 1$). The measures induced by $\{n_k\}_1^N$ under H_0 and H_1 are equivalent for ($0 < \alpha \leqq 2$, $-1 \leqq \beta \leqq 1$); it remains to find the optimum decision rule, the log likelihood ratio, and characterize its performance.

5.1 Likelihood ratio test

Before discussing the general case, the three special analytically tractable cases are treated.

Gaussian ($\alpha = 2$, $-1 \leqq \beta \leqq 1$):

$$p_n(x) = \frac{1}{\sqrt{4\pi}} e^{-x^2/4} \qquad -\infty < x < \infty;$$

$$\therefore \quad l(r_i) = \ln \frac{p_n(r_i/s^1)/s^1}{p_n(r_i/s^0)/s^0} = \ln \frac{s^0}{s^1} - \frac{r_i^2}{4}\left[\left(\frac{1}{s^1}\right)^2 - \left(\frac{1}{s^0}\right)^2\right];$$

$$\therefore \quad \Lambda' = \sum_{i=1}^{N} l(r_i) = N \ln\left(\frac{s^0}{s^1}\right) - \left[\left(\frac{1}{2s^1}\right)^2 - \left(\frac{1}{2s^0}\right)^2\right] \cdot \sum_{i=1}^{N} r_i^2 \underset{H_0}{\overset{H_1}{\gtrless}} L'.$$

The test involves squaring the observations and comparing with a threshold; this test is the well-known chi-squared test (see Ref. 5, pages 163–173).

Cauchy ($\alpha = 1, \beta = 0$):

$$p_n(x) = \frac{1}{\pi}(x^2 + 1)^{-1} \qquad -\infty < x < \infty;$$

$$\therefore \quad l(r_i) = \ln \frac{p_n(r_i/s^1)/s^1}{p_n(r_i/s^0)/s^0} = \ln\left(\frac{s^0}{s^1}\right) - \ln \frac{r_i^2 + (s^0)^2}{r_i^2 + (s^1)^2};$$

$$\therefore \quad \Lambda' = N \ln\left(\frac{s^0}{s^1}\right) - \sum_{i=1}^{N} \ln \frac{r_i^2 + (s^0)^2}{r_i^2 + (s^1)^2} \underset{H_0}{\overset{H_1}{\gtrless}} L'.$$

For $|r_i| \ll s^0, s^1$, Taylor series arguments show $l(r_i)$ behaves as r_i^2, just as in the gaussian case. However, unlike the gaussian case, where $l(r_i)$ behaves asymptotically ($|r_i| \gg s^1, s^0$) as $0(r_i^2)$, here $l(r_i) \sim \ln(s^1/s^0) + 0(r_i^{-2})$; again, large excursions are soft limited, or essentially discarded.

Pearson V ($\alpha = \frac{1}{2}, \beta = -1$):

$$p_n(x) = \begin{cases} \dfrac{1}{\sqrt{2\pi}} x^{-\frac{3}{2}} e^{-\frac{1}{2}x} & x > 0 \\ 0 & x < 0; \end{cases}$$

$$\therefore \quad l(r_i) = \ln \frac{p_n(r_i/s^1)/s^1}{p_n(r_i/s^0)/s^0}$$

$$= \begin{cases} -\dfrac{1}{2}\ln\left(\dfrac{s^0}{s^1}\right) - \dfrac{1}{2r_i}[s_1 - s_0] & (r_i > 0); \\ 0 & (r_i < 0) \end{cases}$$

$$\therefore \quad \Lambda' = -\frac{N}{2}\ln\left(\frac{s^0}{s^1}\right) - \frac{1}{2}(s^1 - s^0)\sum_{i=1}^{N}\left(\frac{1}{r_i}\right) \underset{H_0}{\overset{H_1}{\gtrless}} L'.$$

Again, large deviations in r_i are soft limited or weighted lightly, since asymptotically ($r_i \gg s^1, s^0$)$l(r_i)$ behaves as $0(r_i^{-1})$.

The remaining cases can be treated in identical manner using the power series and asymptotic series expansions for the stable probability density function. For ($0 < \alpha < 2, -1 < \beta < 1; \alpha \neq 1$), the important points are: (*i*) for $|r_i| \ll s^0, s^1$, the ith term ($\beta \neq 0$) in the log likelihood behaves as r_i, unlike in the gaussian case, while for $\beta = 0$, $l(r_i) \sim r_i^2$, (*ii*) for $|r_i| \gg s^0, s^1$, soft limiting of large deviations is used, and the log likelihood's ith term behaves as $\alpha \ln(s^1/s^0) + 0(|r_i|^{-\alpha})$. Figures 13 and 14 show representative log likelihood ratios for fixed α and varying β, and fixed β with α varying, respectively, computed from power series and asymptotic series.[31]

The final case ($0 < \alpha < 2$, $\beta = 1$, or $\beta = -1$) must be handled with a little more care. Only the case $\beta = -1$ is discussed, since the other follows immediately. For ($1 < \alpha < 2$), the first point made above is still valid, while the second point is valid only for $r_i > 0$, $r_i \gg s^0, s^1$.

For $r_i < 0$, $|r_i| \gg s^0, s^1$, $l(r_i)$ behaves as $\alpha \ln (s^1/s^0) + 0(-|r_i|^{1/1-\alpha})$, i.e., decreasing with $|r_i|$. For $(0 < \alpha < 1)$, for $r_i > 0$, $r_i \ll s^0, s^1$, the ith term in the log likelihood behaves as $0(r_i^{-(1/1-\alpha)})$. Finally, for $\alpha = 1$, $l(r_i) = 0\{-\exp[(\pi/2)|r_i|]\}$ as $r_i \to -\infty$.

5.2 Performance limitations

The general prohlem of finding P_E, P_{01}, and P_{10} for arhitrary stable distributions is still open, hoth analytically and numerically (hecause of expense). The three special analytic cases are treated here, to point out the prohlems that must be overcome in the general case, if one attempts to find the log likelihood prohability density hy transform methods.

Gaussian $(\alpha = 2, \ -1 \leqq \beta \leqq 1)$: assuming hypothesis $H_j (j = 0, 1)$ true, the Fourier transform of the log likelihood prohability density is

$$E(e^{iv\Lambda'}|H_1) = \left(\frac{s^0}{s^1}\right)^{ivN} \left\{1 - iv\left[\left(\frac{s^1}{s^0}\right)^2 - 1\right]\right\}^{-N/2}$$

$$E(e^{iv\Lambda'}|H_0) = \left(\frac{s^0}{s^1}\right)^{ivN} \left\{1 + iv\left[\left(\frac{s^0}{s^1}\right)^2 - 1\right]\right\}^{-N/2}.$$

These Fourier transforms can be inverted:

$$\left.\begin{aligned} p(x|H_1) &= \left(\frac{(s^0)^2}{(s^1)^2 - (s^0)^2}\right)^{N/2} x^{(N/2)-1} \\ &\quad \cdot \exp\left(-\frac{(s^0)^2}{(s^1)^2 - (s^0)^2}x\right) \Big/ \Gamma(N/2) \\ p(x|H_0) &= \left(\frac{(s^1)^2}{(s^1)^2 - (s^0)^2}\right)^{N/2} x^{(N/2)-1} \\ &\quad \cdot \exp\left(-\frac{(s^1)^2}{(s^1)^2 - (s^0)^2}x\right) \Big/ \Gamma(N/2) \end{aligned}\right\} x = \Lambda' - N \ln\left(\frac{s^0}{s^1}\right) > 0$$

$$p(x|H_1) = p(x|H_0) = 0 \qquad x = \Lambda' - N \ln\left(\frac{s^0}{s^1}\right) < 0.$$

Finally, the prohahilities of errors of the first and second kind are:

$$\begin{aligned} P_{10} &= 1 - \frac{2}{N}\left(\frac{L'[(s^1)^2 - (s^0)^2]}{(s^0)^2}\right)^{N/2} \\ &\quad \cdot {}_1F_1\left(\frac{N}{2}; 1 + \frac{N}{2}; \frac{L'[(s^0)^2 - (s^1)^2]}{(s^0)^2}\right) \Big/ \Gamma(N/2) \\ P_{01} &= \frac{2}{N}\left(\frac{L'[(s^0)^2 - (s^1)^2]}{(s^1)^2}\right)^{N/2} \\ &\quad \cdot {}_1F_1\left(\frac{N}{2}; 1 + \frac{N}{2}; \frac{L'[(s^0)^2 - (s^1)^2]}{(s^1)^2}\right) \Big/ \Gamma(N/2) \\ &\qquad\qquad L' > N \ln (s^0/s^1) \end{aligned}$$

$$P_{10} = 1, P_{01} = 0 \qquad L' < N \ln (s^0/s^1).$$

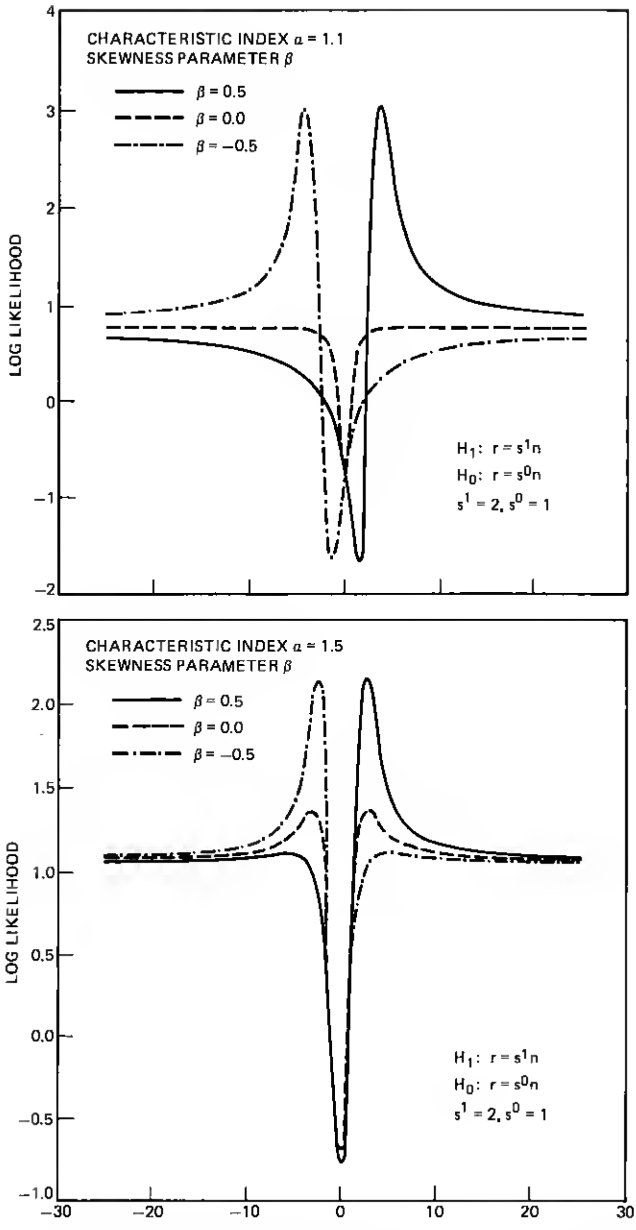

Fig. 13—Representative log likelihood functions (α fixed, β varying).

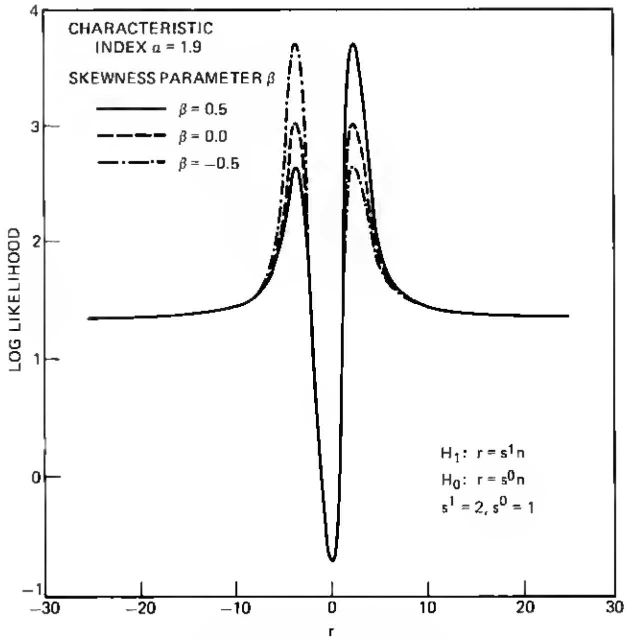

Fig. 13—(continued)

Cauchy ($\alpha = 1, \beta = 0$): the log likelihood characteristic function under H_1 is

$$E(e^{iv\Lambda'}|H_1) = \left\{\int_{-\infty}^{\infty} \frac{s^1}{\pi}\left(\frac{s^1}{s^0}\right)^{iv}[x^2 + (s^0)^2]^{iv}[x^2 + (s^1)^2]^{-iv-1}dx\right\}^N$$
$$= \left\{(s^1/s^0)^{iv+1}{}_2F_1\left[iv + 1, \frac{1}{2}; 1; -1 + \left(\frac{s^1}{s^0}\right)^2\right]\right\}^N,$$

where $s^1 > s^0$ was assumed. Stationary phase arguments show that the characteristic function decays asymptotically as $0(|v|^{-N/2})$. Again, the Fourier inversion lemma guarantees that the problem of finding P_{01} is solved. A similar analysis holds assuming H_0 is true.

An alternate approach is to compute the Mellin transform of the likelihood probability density function; the results are

$$E(\Lambda^{s-1}|H_1) = \left\{\left(\frac{s^1}{s^0}\right)^s {}_2F_1\left[s, \frac{1}{2}; 1; -1 + \left(\frac{s^1}{s^0}\right)^2\right]\right\}^N$$
$$E(\Lambda^{s-1}|H_0) = \left\{\left(\frac{s^1}{s^0}\right)^{s-1} \cdot {}_2F_1\left[s - 1, \frac{1}{2}; 1; -1 + \left(\frac{s^1}{s^0}\right)^2\right]\right\}^N.$$

Unfortunately, it is not clear how to invert this transform to find P_{01} and P_{10}.

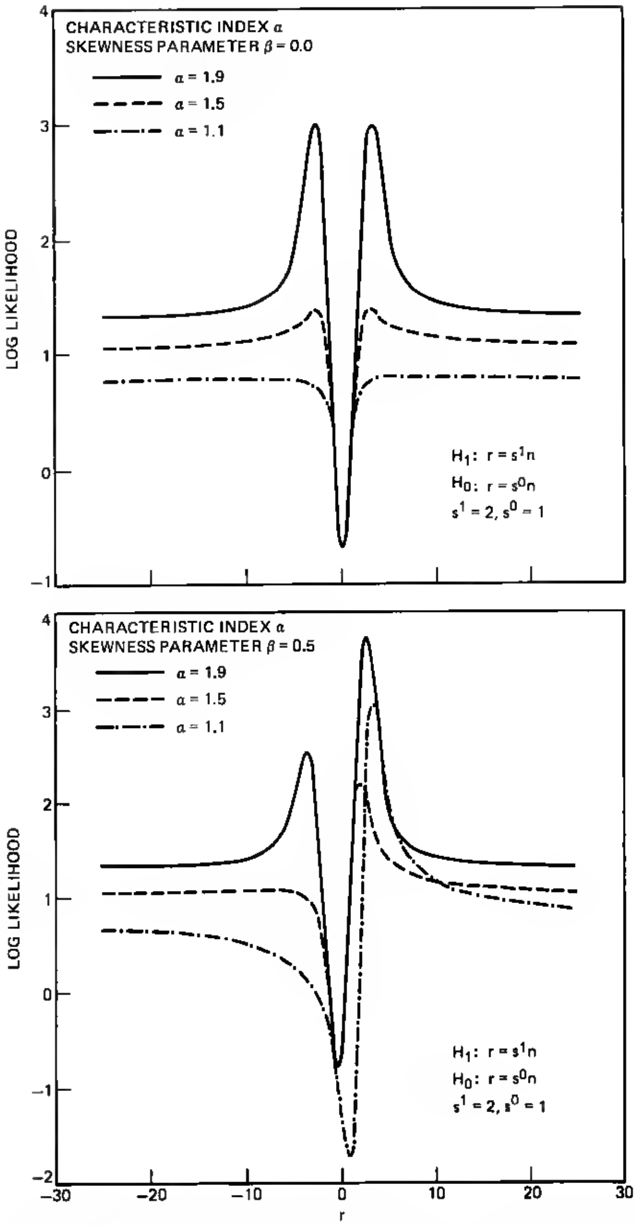

Fig. 14—Representative log likelihood functions (α varying, β fixed).

A third approach is to convolve the probability density of the log likelihood with itself N times; for $N = 2$, the convolution involves elliptic integrals; successive convolutions are quite formidable. This was not investigated further.

Pearson V ($\alpha = \frac{1}{2}, \beta = -1$): the log likelihood characteristic function is (assuming now $s^0 > s^1$)

$$E(e^{iv\Lambda'} | H_1) = \left\{ (s^1)^{iv}(s^0)^{-iv} \Big/ \left[1 - iv \left(\frac{s^0}{s^1} - 1 \right) \right] \right\}^{N/2}$$
$$E(e^{iv\Lambda'} | H_0) = \left\{ (s^1)^{iv}(s^0)^{-iv} \Big/ \left[1 - iv \left(1 - \frac{s^1}{s^0} \right) \right] \right\}^{N/2}.$$

The log likelihood probability density is

$$p(\Lambda' | H_1) = \left(\frac{s^1}{s^0 - s^1} \right)^{N/2} x^{(N/2)-1} e^{-(s^0 - s^1/s^1)x} \Big/ \Gamma\left(\frac{N}{2} \right)$$
$$x = \Lambda' - \frac{N}{2} \ln (s^1/s^0) > 0$$
$$p(\Lambda' | H_0) = \left(\frac{s^0}{s^0 - s^1} \right)^{N/2} x^{(N/2)-1} e^{-(s^0 - s^1/s^0)} \Big/ \Gamma\left(\frac{N}{2} \right)$$
$$p(\Lambda' | H_1) = p(\Lambda' | H_0) = 0 \qquad \Lambda' < \frac{N}{2} \ln (s^1/s^0).$$

The probabilities of errors of the first and second kind are

$$P_{10} = 1 - \frac{2}{N} \left(\frac{L'(s^0 - s^1)}{s^1} \right)^{N/2} {}_1F_1\left(\frac{N}{2}; \frac{N}{2} + 1; \frac{L'(s^1 - s^0)}{s^1} \right) \Big/ \Gamma\left(\frac{N}{2} \right)$$
$$L' > \frac{N}{2} \ln (s^1/s^0)$$
$$P_{01} = \frac{2}{N} \left(\frac{L'(s^0 - s^1)}{s^0} \right)^{N/2} {}_1F_1\left(\frac{N}{2}; \frac{N}{2} + 1; \frac{L'(s^1 - s^0)}{s^0} \right) \Big/ \Gamma\left(\frac{N}{2} \right)$$
$$P_{10} = 1, \qquad P_{01} = 0 \qquad L' < \frac{N}{2} \ln (s^1/s^0).$$

Again, the general problem is still open analytically, because closed-form expressions for stable probability density function are unknown at present (except for the three cases covered here). The general problem is expensive to tackle numerically at present, because of the expense of both calculating and storing the characteristic function of the log likelihood probability density, and because of the expense of repeating these calculations for many different parameter choices.

5.3 Analytic performance bounds

Apparently only in the three special cases does the Kakutani inner product reduce to simple expressions. These results are recorded here, while Section 5.4 discusses numerical approximations of these integrals for various cases of interest.

Gaussian ($\alpha = 2,\ -1 \leqq \beta \leqq 1$):

$$p_n(x) = \frac{1}{\sqrt{4\pi}} e^{-x^2/4} \qquad -\infty < x < \infty$$

$$H_q(P_0, P_1) = \left\{\int_{-\infty}^{\infty}\left[\frac{1}{s^1} p_n\left(\frac{x}{s^1}\right)\right]^q \left[\frac{1}{s^0} p_n\left(\frac{x}{s^0}\right)\right]^{1-q} dx\right\}^N;$$

$$\therefore \quad H_q(P_0, P_1) = (s^1)^{-qN}(s^0)^{-(1-q)N}\left(\frac{q}{(s^1)^2} + \frac{1-q}{(s^0)^2}\right)^{-N/2};$$

$$\therefore \quad H_{\frac{1}{2}}(P_0, P_1) = \left[\frac{1}{2}\left(\frac{s^0}{s^1} + \frac{s^1}{s^0}\right)\right]^{-N/2}.$$

Cauchy ($\alpha = 1,\ \beta = 0$):

$$p_n(x) = \frac{1}{\pi}(x^2+1)^{-1} \qquad -\infty < x < \infty$$

$$H_q(P_0, P_1) = \left\{\int_{-\infty}^{\infty}\left[\frac{1}{s^1} p_n\left(\frac{x}{s^1}\right)\right]^q \left[\frac{1}{s^0} p_n\left(\frac{x}{s^0}\right)\right]^{1-q} dx\right\}^N,$$

where

$$\int_{-\infty}^{\infty}\left[\frac{1}{s^1} p_n\left(\frac{x}{s^1}\right)\right]^q \left[\frac{1}{s^0} p_n\left(\frac{x}{s^0}\right)\right]^{1-q} dx = \left(\frac{s^1}{s^0}\right)^q {}_2F_1\left(\frac{1}{2}, q; 1; -1 + \frac{(s^1)^2}{(s^0)^2}\right).$$

The Hellinger integral can be evaluated from the tables in Ref. 30 263.00:

$$H_{\frac{1}{2}}(P_0, P_1) = \left[\frac{1}{\pi} cn^{-1}\left(-1, \frac{i(s^1 - s^0)}{2\sqrt{s^1 s^0}}\right)^2\right]^N.$$

Pearson V ($\alpha = \frac{1}{2},\ \beta = 1$):

$$p_n(x) = \begin{cases} \dfrac{1}{\sqrt{2\pi}} x^{-\frac{3}{2}} e^{-\frac{1}{2x}} & x \geqq 0 \\ 0 & x < 0; \end{cases}$$

$$\therefore \quad H_q(P_0, P_1) = \left(\frac{(s^1)^q (s^0)^{1-q}}{qs^1 + (1-q)s^0}\right)^{N/2};$$

$$\therefore \quad H_{\frac{1}{2}}(P_0, P_1) = \left(\frac{2\sqrt{s^1 s^0}}{s^1 + s^0}\right)^{N/2}.$$

5.4 Numerical approximation of performance bounds

The methods and checks employed were identical with those used in the detection of location for accurately calculating the inner product of the two stable probability measures.

Figure 15 shows $\mu(q)$ vs q for fixed (s^1/s^0) and $[\alpha = 1.1\ (0.4)\ 1.9, \beta = 0]$. This raises the conjecture that the closer the characteristic index is to 2, the smaller the Kakutani inner product.

Figure 16 shows $\mu(q)$ vs q for fixed (α, β) and various values of (s^1/s^0): the smaller the (s^1/s^0), the smaller the $H_q(P_0, P_1)$.

Figure 17 shows $H_{\frac{1}{2}}(P_0, P_1)$ for various (α, β) as a function of (s^1/s^0); note that the case $\alpha = 2$ does not appear to he singular here.

5.5 Comparison of performance of log likelihood decision rule with a chi-squared test

How does the performance of the log likelihood test compare with that of a chi-squared test, in particular for characteristic index α near 2?

The chi-squared test involves

$$\sum_{i=1}^{N} r_i^2 \underset{H_0}{\overset{H_1}{\gtrless}} L'.$$

The distribution of any one of the r_i^2 can he found from the series described earlier:

$$p(r_i^2 | H_j) = \begin{cases} \dfrac{1}{2s^j\sqrt{r_i}}\, p_n[x = \sqrt{r_i};\, \alpha,\, \beta,\, (s^j)^\alpha,\, \delta = 0] & r_i > 0 \\ 0 & r_i < 0. \end{cases}$$

The discussion now follows from that in Section 4.6, hut is not as detailed. Using elementary arguments (Ref. 15, pages 268–272), it can be shown that if $0 < \alpha < 2$, $-1 < \beta < 1$, then

$$\Pr\left(\sum_{i=1}^{N} r_i^2 > L' | H_j\right) \sim 0(NL'^{-(\alpha/2)}).$$

If L' is set at a threshold which is a fraction of N, then

$$P_E \sim 0(N^{1-(\alpha/2)});$$

i.e., the prohahility of error grows with N, the numher of observations. For comparison, the upper hounds on P_{01}, P_{10}, and P_E for log likelihood detection all hehave as $0(e^{-AN})$, where A depends on $(\alpha, \beta, s^1$, and $s^0)$. Thus, the log likelihood test is asymptotically far superior to the chi-squared hy the ahove argument.

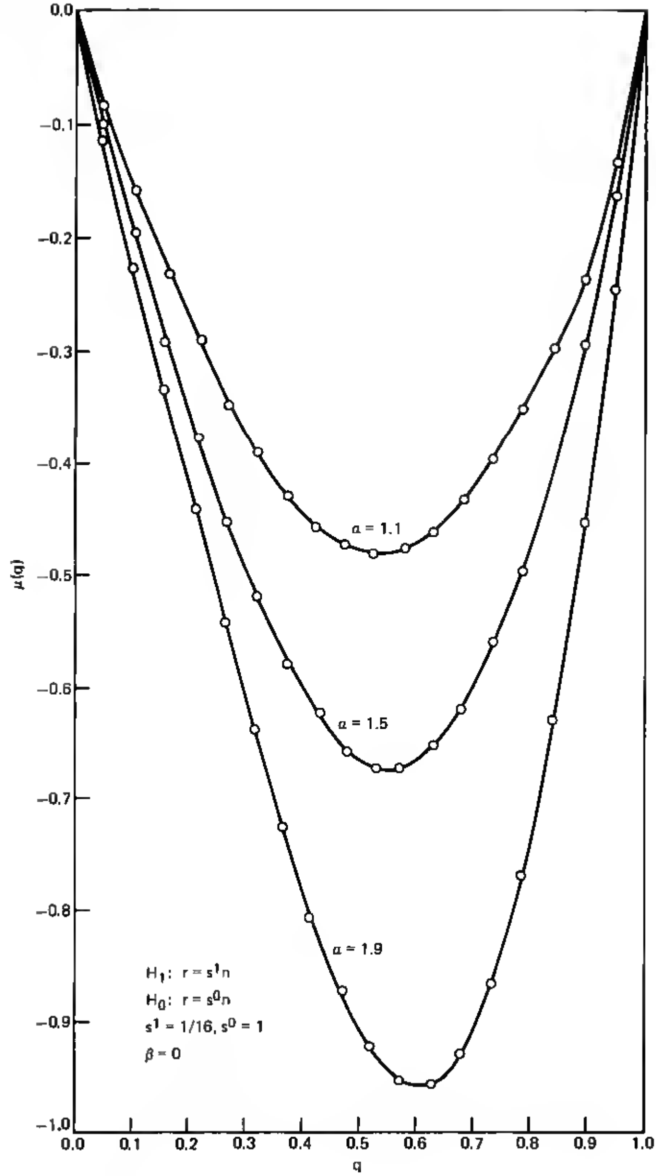

Fig. 15—Logarithm of Kakutani inner product H_q vs q [$\alpha = 1.1(0.4)1.9, \beta = 0$] $(s^0/s^1) = 16$.

VI. DISTINGUISHING STABLE PROBABILITY MEASURES WITH DIFFERENT CHARACTERISTIC INDICES AND SKEWNESS PARAMETERS

For completeness, this section touches on the form the log likelihood test takes for discriminating between stable distributions with different

characteristic indices and with different skewness parameters. Performance of this test will not be covered here; much of the earlier discussion on performance is applicable bere. A table in tbe Appendix summarizes the behavior of $l(r_i)$ both asymptotically and for

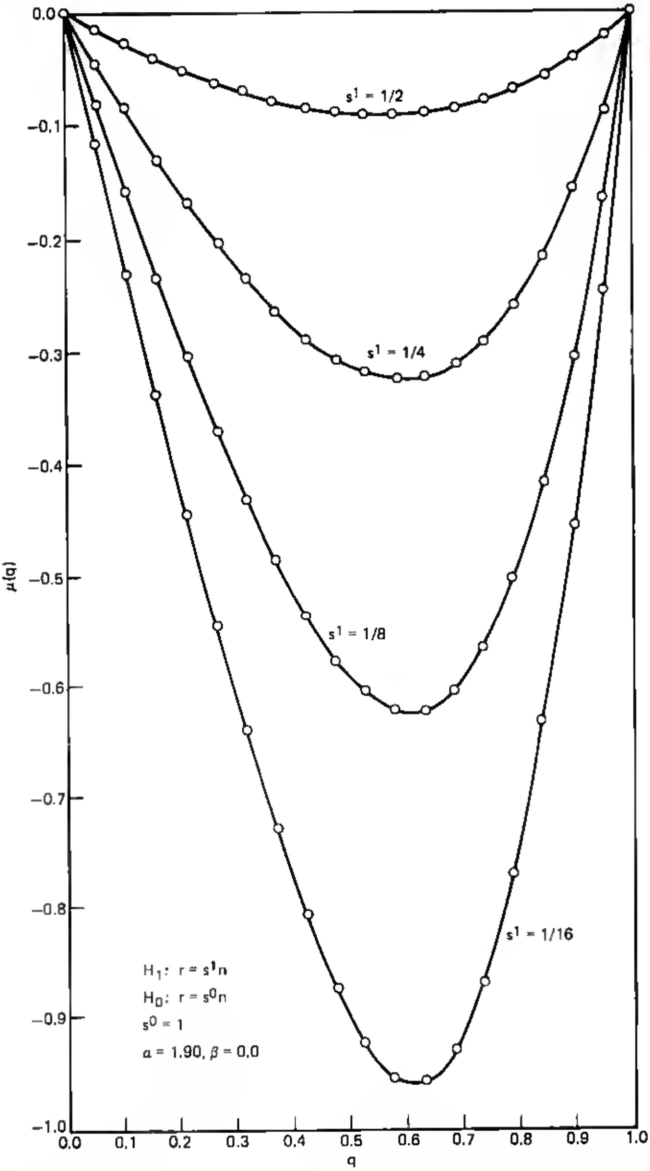

Fig. 16—Logarithm of Kakutani inner product H_q vs q [(α = 1.90, β = 0), (s^0/s^1) = 1, 4, 8, 16].

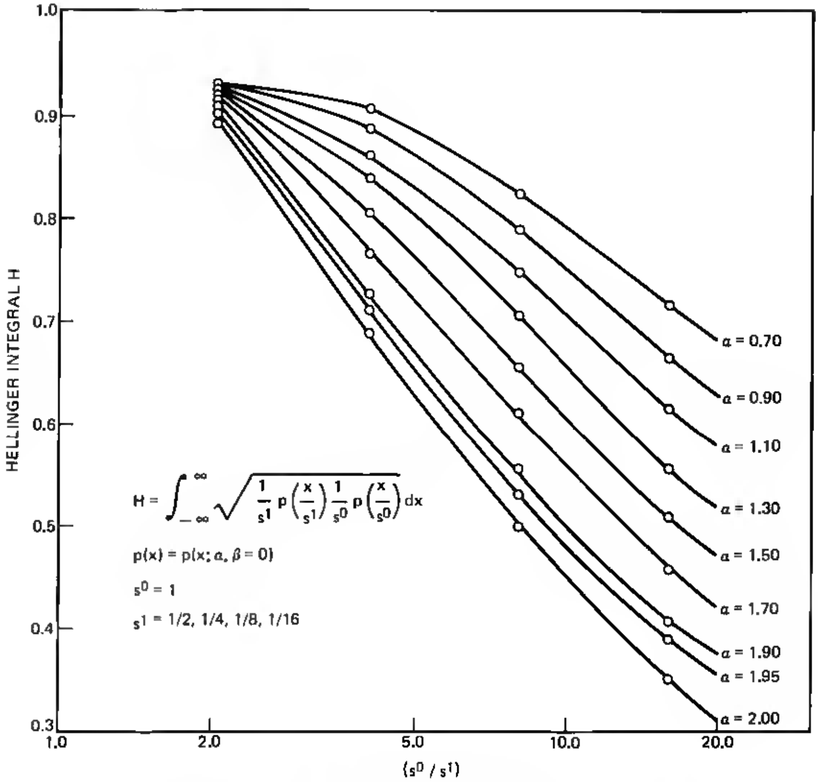

Fig. 17—Hellinger integral vs (s^1/s^0) $[\alpha = 0.7(0.2)1.90, \beta = 0]$.

$|r_i| \ll 1$, and includes both the results in the Sections 5.4 and 5.5 as well as the results of this section.

One of two sequences of i.i.d. stable random variables with known parameters is observed. In Section 6.1, the parameters are $(\alpha^j, \beta, \gamma = 1, \delta = 0)$, where $0 < \alpha^0 < \alpha^1 \leqq 2$; in Section 6.2, the parameters are $(\alpha, \beta^j, \gamma = 1, \delta = 0)$, where $-1 < \beta^0 < \beta^1 \leqq 1$ (recall $j = 0, 1$). The special case $(\alpha = 1, |\beta| = 1)$ is covered in the table in the Appendix but not in the discussion here.

6.1 Distinguishing different characteristic indices

For $-1 < \beta < 1$, the measures P_0 and P_1 are equivalent, so the log likelihood ratio is always finite. The log likelihood test is

$$\Lambda' = \sum_{i=1}^{N} l(r_i) \underset{H_0}{\overset{H_1}{\gtrless}} L',$$

where

$$l(r_i) = \ln \frac{p_n(r_i; \alpha^1, \beta, \gamma = 1, \delta = 0)}{p_n(r_i; \alpha^0, \beta, \gamma = 1, \delta = 0)}.$$

Two cases arise: symmetric ($\beta = 0$) and asymmetric ($\beta \neq 0$, $-1 < \beta < 1$) stable distributions. For the symmetric case, the distributions are symmetric about their unique mode, and thus $l(r_i) \sim |r_i^2|$ for $|r_i| \ll 1$. For the asymmetric case, the modes no longer coincide, and $l(r_i) \sim r_i$ for $|r_i| \ll 1$. Recall that for $1 < \alpha < 2$, for fixed skewness β ($\beta < 0$) the mode decreases as α decreases; for $0 < \alpha < 1$, the opposite is true. Thus, $l(r_i)$ is the difference of two unimodal functions and, in general, should have two points of zero slope. For $|r_i| \gg 1$, $l(r_i) = 0(-\ln |r_i|)$, so large deviations are weighted quite strongly. Note the log likelihood distribution has its support on whole line, unlike the two previous sections, except for $(0 < \alpha_0 < 1 \leqq \alpha_1 \leqq 2$, $|\beta| = 1)$.

For $\beta = -1$, and $1 < \alpha^0 < \alpha^1 < 2$, the measures P_0 and P_1 are equivalent, and the above discussion follows immediately with one exception: for $r_i \gg 1$, $l(r_i) = 0(-\ln r_i)$, while for $|r_i| \gg 1$, $r_i < 0$, $l(r_i) = 0(|r_i|^{\alpha_0/\alpha_0 - 1})$.

For $\beta = -1$, $0 < \alpha^0 < \alpha^1 < 1$, the measures P_0 and P_1 are equivalent. For $r_i > 0$, $|r_i| \ll 1$, $l(r_i) \sim r_i^{\alpha_0/1-\alpha_0}$, while for $r_i \gg 1$, $l(r_i) = 0(-\ln r_i)$.

Finally, for $\beta = -1$, $0 < \alpha^0 < 1 < \alpha^1 < 2$, the measures P_0 and P_1 are neither equivalent nor mutually orthogonal. For $r_i \gg 1$, $l(r_i) = 0(-\ln r_i)$, while for $r_i < 0$, $l(r_i) = \infty$. For $r_i > 0$, $r_i \ll 1$, $l(r_i) = 0(r_i^{\alpha_0/\alpha_0 - 1})$.

6.2 *Distinguishing different skewness parameters*

For $-1 < \beta^0 < \beta^1 < 1$, the measures P_0 and P_1 are equivalent, so the log likelihood ratio is finite. The discussion follows that of Section 6.1 exactly, with the difference that if $r_i \gg 1$, $l(r_i) = \ln (R_1/R_0) + 0(r_i^{-\alpha})$, while if $|r_i| \gg 1$, $r_i < 0$, $l(r_i) = \ln (L_1/L_0) + 0(|r_i|^{-\alpha})$.*

For $-1 = \beta^0 < \beta^1 < 1$, $1 \leqq \alpha < 2$, the measures P_0 and P_1 are equivalent. For $|r_i| \gg 1$, $r_i < 0$, $l(r_i) = 0(|r_i|^{\alpha/\alpha-1})$, while for $r_i \gg 1$, $l(r_i) = \ln (R_1/R_0) + 0(r_i^{-\alpha})$.

For $-1 = \beta^0 < \beta^1 < 1$, $0 < \alpha < 1$, the measures P_0 and P_1 are neither equivalent nor mutually orthogonal. For $r_i \gg 1$, $l(r_i) = \ln (R_1/R_0) + 0(r_i^{-\alpha})$, while for $0 < r_i \ll 1$, $l(r_i) = 0(r_i^{\alpha/\alpha-1})$. For $0 > r_i$, $l(r_i) = \infty$.

VII. ACKNOWLEDGMENTS

S. O. Rice and J. A. Morrison were quite helpful in evaluating the performance limitations of the likelihood ratio test for distinguishing between one of two location parameters. J. L. Blue had many suggestions concerning the numerical analysis. M.-L. Thomson provided a useful computer program for calculating hypergeometric functions.

* See the Appendix for definition of constants R_j, $L_j (j = 0, 1)$.

APPENDIX

Asymptotic Behavior of Log Likelihood Ratio

A.1 Location (δ)

$$l(x) = \ln \frac{p(x;\alpha,\beta,\gamma,\delta_1)}{p(x;\alpha,\beta,\gamma,\delta_0)} \qquad \delta_0 < \delta_1$$

	$x \to +\infty$	$x \to -\infty$
$\alpha = 2$	$0(x)$	$0(x)$
$0 < \alpha < 2, -1 < \beta < 1$	$0(x^{-1})$	$0(x^{-1})$
$1 < \alpha < 2, \beta = -1$	$0(x^{-1})$	$0(-\lvert x\rvert^{1/\alpha-1})$
$\alpha = 1, \beta = -1$	$0(x^{-1})$	$0(-e^{(\pi/2)\lvert x-\delta_1\rvert})$
	$x \to +\infty$	$x \downarrow \delta_1$
$0 < \alpha < 1, \beta = -1$	$0(x^{-1})$	$0(-(x-\delta_1)^{\alpha/\alpha-1})$.

A.2 Scale (c)

$$l(x) = \ln \frac{p(x;\alpha,\beta,\gamma_1 = c_1^\alpha,\delta = 0)}{p(x;\alpha,\beta,\gamma_0 = c_0^\alpha,\delta = 0)} \qquad c_0 < c_1$$

	$x \to +\infty$	$x \to -\infty$
$\alpha = 2$	$0(x^2)$	$0(x^2)$
$0 < \alpha < 2, -1 < \beta < 1$*	$\alpha \ln (c_1/c_0) + 0(x^{-\alpha})$	$\alpha \ln (c_1/c_0) + 0(\lvert x\rvert^{-\alpha})$
$1 < \alpha < 2, \beta = -1$	$\alpha \ln (c_1/c_0) + 0(x^{-\alpha})$	$0(-\lvert x\rvert^{\alpha/\alpha-1})$
$\alpha = 1, \beta = -1$	$\alpha \ln (c_1/c_0) + 0(x^{-\alpha})$	$0(-e^{(\pi/2)\lvert x/c_1\rvert})$
	$x \to +\infty$	$x \downarrow 0$
$0 < \alpha < 1, \beta = -1$	$\alpha \ln (c_1/c_0) + 0(x^{-\alpha})$	$0(-x^{\alpha/\alpha-1})$.

A.3 Characteristic index (α)

$$l(x) = \ln \frac{p(x;\alpha_1,\beta,\gamma = 1,\delta = 0)}{p(x;\alpha_0,\beta,\gamma = 1,\delta = 0)}, \qquad 0 < \alpha_0 < \alpha_1 \leqq 2$$

	$x \to +\infty$	$x \to -\infty$
$0 < \alpha_0 < \alpha_1 = 2, -1 < \beta < 1$	$0(-x^2)$	$0(-x^2)$
$0 < \alpha_0 < \alpha_1 < 2, -1 < \beta < 1$	$0(-\ln x)$	$0(-\ln \lvert x\rvert)$
$1 < \alpha_0 < \alpha_1 = 2, \beta = -1$	$0(-x^2)$	$0(\lvert x\rvert^{\alpha_0/\alpha_0-1})$
$1 < \alpha_0 < \alpha_1 < 2, \beta = -1$	$0(-\ln x)$	$0(\lvert x\rvert^{\alpha_0/\alpha_0-1})$
$1 = \alpha_0 < \alpha_1 = 2, \beta = -1$	$0(-x^2)$	$0(e^{(\pi/2)\lvert x\rvert})$
$1 = \alpha_0 < \alpha_1 < 2, \beta = -1$	$0(-\ln x)$	$0(e^{(\pi/2)\lvert x\rvert})$
	$x \to +\infty$	$x \downarrow 0$
$0 < \alpha_0 < 1 < \alpha_1 < 2, \beta = -1$	$0(-\ln x)$	$0(x^{\alpha_0/\alpha_0-1})$
$0 < \alpha_0 < \alpha_1 = 1, \beta = -1$	$0(-\ln x)$	$0(x^{\alpha_0/\alpha_0-1})$
$0 < \alpha_0 < \alpha_1 < 1, \beta = -1$	$0(-\ln x)$	$0(x^{\alpha_0/\alpha_0-1})$.

* This excludes the Cauchy ($\alpha = 1, \beta = 0$), which was examined in the text as a special case.

A.4 Skewness (β)

$$l(x) = \ln \frac{p(x;\alpha,\beta_1,\gamma=1,\delta=0)}{p(x;\alpha,\beta_0,\gamma=1,\delta=0)} \qquad -1 \leqq \beta_0 < \beta_1 \leqq 1$$

	$x \to +\infty$	$x \to -\infty$
$0 < \alpha < 2,\ -1 < \beta_0 < \beta_1 < 1$	$\ln (R_1/R_0) + 0(x^{-\alpha})$	$\ln (L_1/L_0) + 0(\lvert x\rvert^{-\alpha})$
$1 < \alpha < 2,\ -1 = \beta_0 < \beta_1 < 1$	$\ln (R_1/R_0) + 0(x^{-\alpha})$	$0(\lvert x\rvert^{\alpha/\alpha-1})$
$\alpha = 1,\ -1 = \beta_0 < \beta_1 < 1$	$\ln (R_1/R_0) + 0(x^{-\alpha})$	$0(e^{(\pi/2)\lvert x\rvert})$
$1 < \alpha < 2,\ -1 = \beta_0,\ 1 = \beta_1$	$0(-x^{\alpha/\alpha-1})$	$0(\lvert x\rvert^{\alpha/\alpha-1})$
$\alpha = 1,\ -1 = \beta_0,\ 1 = \beta_1$	$0(-e^{(\pi/2)\lvert x\rvert})$	$0(e^{(\pi/2)\lvert x\rvert})$

	$x \to +\infty$	$x \downarrow 0$
$0 < \alpha < 1,\ -1 = \beta_0 < \beta_1 < 1$	$\ln (R_1/R_0) + 0(x^{-\alpha})$	$0(x^{\alpha/\alpha-1})$

$$R_j = \sin \frac{\pi}{2}(\theta_j - \alpha),\ \tan(\pi\theta_j/2) = \beta_j \tan(\pi\alpha/2)$$

$$j = 0, 1$$

$$L_j = \sin \frac{\pi}{2}(\theta_j - \alpha),\ \tan(\pi\theta_j/2) = -\beta_j \tan(\pi\alpha/2).$$

REFERENCES

1. B. W. Stuck and B. Kleiner, "A Statistical Analysis of Telephone Noise," B.S.T.J., *53* (September 1974), pp. 1263–1320.
2. J. E. Evans, "Preliminary Analysis of ELF Noise," M.I.T. Lincoln Laboratory Tech. Note 1969-18 (26 March 1969), DDC AD691814.
3. P. Cootner ed., *The Random Character of Stock Market Prices*, Cambridge, Mass.: M.I.T. Press, 1964.
4. E. F. Fama, "The Behavior of Stock Market Prices," J. Bus., *38* (January 1965), pp. 34–105.
5. E. H. Lehmann, *Testing Statistical Hypotheses*, New York: Wiley, 1959.
6. P. Huber, "A Robust Version of the Probability Ratio Test," Ann. Math. Statist., *36* (December 1965), pp. 1753–1758.
7. R. M. Lerner and V. R. Algazi, "Binary Detection in White Non-Gaussian Noise," Available from University of Calif. (Davis) Elec. Eng. Dept.
8. J. E. Evans and A. S. Griffiths, "Design of a Sanguine Noise Processor Based Upon World-Wide Extremely Low Frequency (ELF) Recordings," IEEE Trans. Commun. Theory, *22* (April 1974), pp. 528–539.
9. J. Neveu, *Mathematical Foundations of the Calculus of Probability*, San Francisco; Holden-Day, 1965.
10. J. F. C. Kingman and S. J. Taylor, *Introduction to Measure and Probability*, Cambridge University Press, 1965.
11. J. L. Doob, *Stochastic Processes*, New York: Wiley, 1953.
12. L. Breiman, *Probability*, Reading, Mass.: Addison-Wesley, 1968.
13. B. V. Gnedenko and A. N. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables*, Revised, Reading, Mass.: Addison-Wesley, 1968.
14. E. Lukacs, *Characteristic Functions*, 2nd ed., N. Y.: Hafner, 1968.
15. W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. II, N. Y.: Wiley, 1966.
16. M. Lipschutz, "On Strong Bounds for Sums of Independent Random Variables Which Tend to a Stable Distribution," Trans. Amer. Math. Soc., *81* (January–April, 1956), pp. 135–154.
17. I. A. Ibragimov and Yu. V. Linnik, *Independent and Stationary Sequences of Random Variables*, Groningen, Netherlands: Wolters-Nordhoff, 1971, pp. 47–76.

18. A. V. Skorokhod, "Asymptotic Formulas for Stable Distribution Laws," Transl. Math. Statist. Probl., *1*, Providence, R. I.: American Matbematical Society, 1961, pp. 157–161.
19. P. A. Frost, "Some Singular Estimation and Detection Problems for Conditionally Independent Increment Processes," p. 232, Proc. Ninth Annual Allerton Conf., 1971, University of Illinois.
20. M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, New York: Dover, 1958.
21. L. J. Slater, *Confluent Hypergeometric Functions*, Cambridge University Press, 1965.
22. H. Dym and H. P. McKean, *Fourier Series and Integrals*, New York: Academic, 1972.
23. E. T. Copson, *Asymptotic Expansions*, Cambridge University Press, 1965.
24. E. J. Brody, "An Elementary Proof of the Gaussian Dichotomy Theorem," Z. Wahrschein. Verwandte Geb., *20* (November 1971), pp. 217–226.
25. C. E. Shannon, R. G. Gallager, and E. R. Berlekamp, "Lower Bounds to Error Probabilities for Coding on Discrete Memoryless Cbannels, I," Inform. Contr., *10* (January 1967), pp. 65–103.
26. H. Chernoff, "A Measure of Asymptotic Efficiency for Tests of a Hypothesis Based on a Sum of Observations," Ann. Math. Statist., *23* (December 1952), pp. 493–507.
27. C. H. Kraft, "Some Conditions for Consistency and Uniform Consistency of Statistical Procedures," University of Calif. (Berkeley), Publ. Statistics *2*, No. 6 (1955), pp. 125–142.
28. T. T. Kadota and L. A. Sbepp, "On the Best Set of Linear Observables for Discriminating Two Gaussian Signals," IEEE Trans. Inform. Theory, *13* (April 1967), pp. 278–284.
29. M. E. Hellman and J. Raviv, "Probability of Error, Equivocation, and the Chernoff Bound," IEEE Trans. Inform. Theory, *16* (July 1970), pp. 368–372.
30. P. F. Byrd and M. D. Friedman, *Handbook of Elliptic Integrals for Engineers and Physicists*, 2nd ed., New York: Springer, 1971.
31. M.-J. Cross, "Tables of Finite Mean Nonsymmetric Stable Distributions as Computed From Their Convergent and Asymptotic Series," J. Statist. Computat. Simul., *3* (January 1974), pp. 1–27.
32. V. M. Zolotarev, "On Representation of Stable Laws by Integrals," Transl. Matb. Statist. Prob., *6*, Providence, R. I.: American Mathematical Society, 1966, pp. 84–88.
33. W. H. DuMouchel, "Stable Distributions in Statistical Inference," Yale University, Ph.D. dissertation, 1971.
34. J. Cooley, P. Lewis, and P. Welch, "Tbe Fast Fourier Transform Algoritbm and Its Applications," IBM Research Paper RC-1743, 9 Feb. 1967.
35. F. L. Bauer, H. Rutishauser, and E. Stiefel, "New Aspects in Numerical Quadrature," pp. 199–218, Proc. Symp. App. Math., *15*, Am. Math. Soc., Providence, R. I., 1963.